University of São Paulo
College of Economics, Business and Accounting (FEA/USP)
Dept. of Accountancy and Actuarial Science

Marcelo dos Santos Guzella

Investor Attention in the Brazilian Stock Market: Essays in Behavioral Finance

Atenção do Investidor no Mercado Brasileiro de Ações: Ensaios em Finanças Comportamentais

São Paulo
2020

Prof. DSc. Vahan Agopyan
Dean of the University of São Paulo

Prof. DSc. Fábio Frezatti
Director of the College of Economics, Business and Accounting (FEA/USP)

Prof. DSc. Valmor Slomski
Chief Officer of the Dept. of Accountancy and Actuarial Science

Prof. DSc. Lucas Ayres Barreira de Campos Barros
Coordinator of the Postgraduate Program in Controllership and Accountancy

Marcelo dos Santos Guzella

Investor Attention in the Brazilian Stock Market: Essays in Behavioral Finance

Dissertation presented to the Dept. of Accountancy and Actuarial Science of the College of Economics, Business and Accounting (FEA/USP) of the University of São Paulo as partial requirement for obtaining the title of Doctor of Science.

Supervisor: Prof. DSc. Francisco Henrique Figueiredo de Castro Junior

Final version

São Paulo
2020

Catalog Card
Prepared by the Dept. of Technical Processing of SBD/FEA/USP

Guzella, Marcelo dos Santos.
Investor Attention in the Brazilian Stock Market: Essays in Behavioral Finance / Marcelo dos Santos Guzella. – São Paulo, 2020.
114 p.

Dissertation (Doctoral Program) – University of São Paulo, 2020.
Supervisor: Francisco Henrique Figueiredo de Castro Junior.

1. Investor Attention. 2. Behavioral Finance. 3. Internet Search Volume. 4. Stock Markets. 5. Volatility. I. University of São Paulo. College of Economics, Business and Accounting (FEA/USP). II. Title.

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel (CAPES) - Finance Code 001.

To my family.

Acknowledgements

It is an impossible task to mention exhaustively all those who have somehow contributed to this meaningful accomplishment.

First I would like to thank God for the opportunity, strength and ability to undertake this research and to conclude it reasonably.

I also wish to express my deepest gratitude to my family for encouraging and supporting me during this work: my mother, Neusa, and particularly my father, Fernando, who while alive did not save efforts to encourage me to pursue my life aspirations; and my brothers, Thiago, Matheus and Rodrigo, whose academic achievements inspired me to carry out this challenging mission. I will not run the risk of mentioning other family members and friends who helped me on this journey, but here is my appreciation and consideration to all those ones whose assistance was a milestone in the completion of this project.

I would like to pay my special regards to my Supervisor, Professor Henrique Castro, for the valuable ideas and assistance along the research process. Your great advice proved monumental towards the success of this study. I also thank Professor Veronica Santana for always being so open and patient while advising me on so many methodological decisions and issues.

I also recognize the invaluable support of the Department Coordinator, Professor Lucas Barros, and all the other professors of the Department, particularly the ones with whom I had the privilege to have classes throughout the Doctoral Program. Moreover, I cannot forget the important support of the Department's technical and support staff, as well as my doctoral colleagues.

The experience at Columbia Business School was also of great value for this academic achievement, with special thanks to my Visiting Supervisor Professor Kent Daniel and Professors Xavier Giroud and Stijn Van Nieuwerburgh.

The professors of Coppead UFRJ and EM Lyon also deserve sincere thanks. Those are distinguished Business Schools where I studied during the Master's Program and got academic and methodological bases that were fundamental for conducting this research.

I wish to show my gratitude to the Professors that, together with my Supervisor, composed the Qualification Board, Lucas Barros, Ricardo Rochman and Andrea Minardi, as well as the the participants of the Brazilian Conference of Behavioral Finance and the Brazilian Finance Meeting, who gave rich comments and suggestions to improve this research endeavor.

Last but not least, I am indebted to Codemge (Development Company of Minas Gerais) for the support needed to concile professional activities with the Doctoral Program and this dissertation, so I express my gratitude to former and current directors (specially Mr. Ricardo Toledo and Mr. Marco Antonio Castello Branco) and colleagues.

Abstract

Guzella, M. S. (2020). *Investor Attention in the Brazilian Stock Market: Essays in Behavioral Finance* (PhD Dissertation, University of São Paulo, São Paulo).

We developed three essays regarding the impact of the investor attention on the Brazilian stock market. Attention is a cognitive resource of great relevance and has been increasingly studied in research related to behavioral finance. It has a crucial role in processes such as buying and selling assets, absorption of information and risk management. Firstly, we evidenced that attention transmits market efficiency for being a requirement for the discovery of released information, and this effect is more associated with professional attention. After that, we verified that attention, particularly the one of retail investors, is capable of inducing volatility to the market due to a price pressure by noise trading. Finally, we verified that the volatility of prices is less asymmetric when investors are more attentive to financial information. We measured attention through the volume of searches for financial information over the Internet, particularly the queries performed using Google and Bloomberg. These indicators have properties that allow several approaches that were limited or impossible before they were available. All the results are robust to different methodologies and specifications. Among other innovations, this study is a pioneer in isolating the effect of retail and professional attention on the market efficiency and asymmetry. Our findings contribute to a better understanding of the influence of aggregate psychological aspects on stock prices and open promising venues for research ideas in many fields.

Keywords: Investor Attention, Behavioral Finance, Internet Search Volume, Stock Markets, Volatility.

Contents

Acronyms **7**

1 Introduction **9**
1.1 Attention Variables 13
1.1.1 Retail Investor Attention: Google Search Volume (GSV_t) 13
1.1.2 Professional Investor Attention: Bloomberg Search Volume (BSV_t) 14
1.1.3 Isolated Non-Professional Attention: Google Search Volume Residuals ($rGSV_t$) 15
References 16

2 Information Discovery: Effects of Investor Attention on the Market Efficiency **19**
2.1 Literature Review 21
2.1.1 Information Discovery Hypothesis 21
2.1.2 Competition for Attention 26
2.1.3 Attention and Market Efficieny 28
2.2 Methodology 30
2.2.1 Data Description 31
2.2.2 Modeling 32
2.3 Empirical Results 34
2.3.1 Descriptive Analysis 34
2.3.2 Autoregressive Modeling 37
2.3.3 Effects of Filtered Non-Professional Attention 40
2.4 Final Considerations of the Chapter 41
References 43

3 Price Pressure Induction: Effects of the Investor Attention on the Market Volatility **47**
3.1 Literature Review 48
3.1.1 Behavioral Biases in Financial Markets 48
3.1.2 Volatility Induction by Noise Traders 49
3.1.3 Effects of the Attention Level on Financial Markets 50
3.1.4 Internet Search Queries as an Indicator of Investor Attention 52
3.1.5 Empirical Evidence about Investor Attention in Brazil 56
3.1.6 Additional Considerations Regarding Attention and Financial Markets 57
3.2 Methodology 58
3.2.1 Data Description 58
3.2.2 Modeling 60

3.3 Empirical Results . . . 61
3.3.1 Descriptive Analysis . . . 61
3.3.2 Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) Modeling . . . 62
3.3.3 Vector Autoregressive (VAR) Modeling . . . 65
3.4 Final Considerations of the Chapter . . . 68
References . . . 70

4 Ostrich Behavior: Effects of Investor Attention on the Volatility Asymmetry 73
4.1 Literature Review . . . 76
4.1.1 Volatility Asymmetry . . . 77
4.1.2 Attention Reaction to the Asset Prices . . . 81
4.2 Methodology . . . 87
4.2.1 Data Description . . . 87
4.2.2 Variance Asymmetry Modeling . . . 89
4.2.3 Modeling the Relationship between Asymmetry and Attention . . . 91
4.3 Empirical Results . . . 91
4.3.1 Variance Asymmetry Modeling Outcomes . . . 92
4.3.2 Outcomes of the Relationship between Asymmetry and Attention . 94
4.3.3 Isolating the Effects of Professional and Retail Attention on Asymmetry . . . 97
4.4 Final Considerations of the Chapter . . . 98
References . . . 100

5 Concluding Remarks 105

References 107

List of Tables

2.1 Sample of stocks 33
2.2 Descriptive statistics of the time series 35
2.3 Autoregressive estimation for AmBev (largest Brazilian beverage company) 38
2.4 Autoregressive estimation for CSN (large Brazilian steel company) 39
2.5 Summary of the results for the stocks 40
2.6 Results using $rGSV$ (non-professional attention uncorrelated with professional one) 42

3.1 Correlation between "Bovespa Index (Ibovespa)"-related keywords 62
3.2 Descriptive statistics of the time series 62
3.3 GARCH modeling of the Ibovespa log-returns 64
3.4 Correlation matrix of the series 65
3.5 Vector autoregressive modeling 67

4.1 Descriptive statistics 89
4.2 Asymmetry Modeling of the Ibovespa 93
4.3 Parameters estimation of the longitudinal determinants of Ibovespa daily volatility asymmetry 95
4.4 Average asymmetry in each quartile of returns and attention 96
4.5 Estimates of the effect of professional and retail attention on the asymmetry of the Ibovespa daily volatility 99

List of Figures

1.1 Search engines market share in Brazil 12
1.2 Google Trends 13
1.3 Bloomberg share of users by job title and Bloomberg terminal 15

2.1 Flow diagram of the role of attention in the incorporation of information into prices 30
2.2 Frequency distribution of search volume correlations among stocks 36
2.3 Frequency distribution of Bloomberg and Google search volume correlations 37

3.1 Flow diagram of the role of non-professional attention in the volatility of prices 59
3.2 Graphical visualization of the series 63
3.3 Impulse response functions of the series 69

4.1 Illustrative model of the relationship between attention and volatility asymmetry 86
4.2 Frequency distribution of the asymmetry parameter of the Ibovespa daily volatility 94

Acronyms

AIC Akaike Information Criteria. 61, 64, 65, 93
Amex American Stock Exchange. 51
APARCH Asymmetric Power Autoregressive Conditional Heteroskedasticity. 75, 76, 89, 90, 92, 93, 94, 95, 96, 98, 99
ARCH AutoRegressive Conditional Heteroskedasticity. 60, 63, 79, 90, 93
ARMA Autoregressive–Moving-Average. 89, 90, 92, 93, 94, 95, 96, 99

B3 Brasil Bolsa Balcão S/A. 33, 61, 87, 88
BEKK Baba-Engle-Kraft-Kroner. 79
BIC Bayesian Information Criteria. 61, 64, 65, 93

CAPM Capital Asset Pricing Model. 56, 77, 78, 79
CVM Brazilian Securities Exchange Commission. 10

Dow Dow Jones Industrial Average. 28

FELM Fixed Effects Linear Model. 94, 95, 98, 99
ForEx Foreign Exchange. 70

GARCH Generalized AutoRegressive Conditional Heteroskedasticity. 2, 3, 60, 62, 63, 64, 65, 67, 69, 79, 89, 90
GLS Generalized Least Squares. 91, 94, 95, 98, 99
GMM Generalized Method of Moments. 56

HAR Heterogeneous Autoregressive. 54, 55

Ibovespa Bovespa Index. 3, 5, 31, 33, 35, 36, 37, 42, 56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 87, 88, 89, 91, 92, 93, 94, 95, 96, 97, 99
IPO Initial Public Offering. 25, 53, 70

LM test Lagrange Multiplier test. 61, 63, 64

Nasdaq National Association of Securities Dealers Automated Quotations. 28, 51
NYSE New York Stock Exchange. 24, 51

PCA Principal Component Analysis. 88, 97

US United States. 55, 66, 68, 74, 75, 81, 92, 96, 100

VAR Vector Autoregressive. 2, 52, 54, 56, 65

1 Introduction

> As I grow older, I pay less attention to what men say. I just watch what they do.
>
> Andrew Carnegie (1835-1919)
> Industrialist, philanthropist

Being attentive to something is concentrating mentally on a task or target. Attention implies selection, since there are always alternative activities or things to be engaged with. Besides selection, it is a matter of intensity as well. As Kahneman (1973) mentions, a student that reads a comic book, while his professor is teaching, should be blamed for improperly selecting what to be attentive to. On the other side, the student that is immersed in a pleasant state of drowsiness not only fails on focusing his attention, but also has less amount of attention to be focused.

Attention is then necessary condition to full absorption of the available information. On the other side, the amount of information globally generated and stored is increasingly abundant, particularly in digital media. Gillings et al. (2016) estimate that, during the last three decades, the amount of digital information stored has doubled each 2.5 years, approximately. The researchers calculate that this storage may have reached 5 *zettabytes* (5×10^{21} *bytes*) in 2014, which is 500 times the amount of information in the genome of all of the 7.2 billion inhabitants on Earth (considering the space in bytes after coding it).

Given this current perspective, there are no doubts that we live in the Information Era, dominated by an increasing integration of the markets and a constant need of agility in decisions. Excess of information consumes too much attention, which is a scarce resource (Hirshleifer, 2001; Kahneman, 1973). The immediate incorporation of information on prices requires efficient allocation of attention, which may not occur. Attention helps to restrain the bunch of options (and minimize search costs), but does not always have a straight relationship with utility. This reality justifies an analysis of the behavior of the agents under this approach and its implications to asset prices.

In this dissertation, agents are considered those entities with some participation or influence on markets. In the stock market, institutional and individual investors are among the most relevant agents, as well as regulation agents, the government and companies that perform the transactions or whose stocks are traded there. Additionally, we use the terms individual, retail, non-sophisticated and non-professional investors interchangeably. The same applies for professional, sophisticated and institutional investors. Those terms do not necessarily match the definitions used by the Brazilian Securities Exchange Commission (CVM) for formal and legal purposes.

Individuals use a variety of heuristics to estimate the chances of occurrence of events. These mental shortcuts can generate biases that result in consistently wrong probabilities assessments, as confirmed in many experiments (Tversky & Kahneman, 1975). Since these biases have the potential to make them believe in trends and patterns (that indeed do not exist) in time series of stock prices (Andreassen & Kraus, 1990), this makes them strong candidates to noise traders. This class of investor does not act based on fundamental values and tends to react exaggeratedly to good and bad news (De Long et al., 1990). Burton and Shah (2013) stated that they do not trade on true information, but on aged and phony information. Many of them trade speculatively and based on chart analysis, aiming for short-term capital gains. Nonetheless, their participation in stock exchanges contributes significantly to market liquidity (Kaniel et al., 2008).

Professional investors, on the other hand, generally have a significant share of the market, which gives them an also relevant role in the dynamics of asset prices (Ben-Rephael et al., 2017). This perspective — and the fact that the literature focusing on the attention of this class is scarce — indicates the importance of an approach that takes into account their isolated influence.

Measures that represent the attention of the market participants have been used frequently to a more comprehensive understanding of attributes such as bounded rationality and heterogeneity of beliefs, and their effects. Related literature indicated that these measures permit the identification or anticipation of the turbulence level of the capital market (Barber & Odean, 2008; Dimpfl & Jank, 2016; Huddart et al., 2009), of the general performance of the economy, or even of specific macroeconomic variables, such as unemployment rate, inflation or industrial production (Gomes & Taamouti, 2016).

Since attention is a non-observable measure, researchers and practitioners adopt proxies that may capture part of the variance, such as trading volume, past abnormal returns or presence in media news (Barber & Odean, 2008; Da et al., 2011; Huddart et al., 2009).

Besides, by affecting economic dynamics, attention measures may have a relevant relationship with comovement changes of different asset classes, or of cross-country markets (Gomes & Taamouti, 2016).

However, variables such as trading volume and returns, for being results of equilibrium, may not reasonably satisfy the role of attention measure. The presence in the news, by itself, also do not guarantee that attention was captured by the agents (Da et al., 2011).

The amount of search queries of specific keywords in popular engines can be a direct and effective proxy for the level of attention of non-professional investors. This type of agent has less access to sophisticated platforms and typically focus attention only when they need to obtain information about the market. Da et al. (2011) pointed that this is an unquestionable measure of revealed attention, since, when we search about a topic, we are automatically attentive to it. The potential provided by this measure was already explored to anticipate the activity of a wide range of markets, such as real estate, automotive and tourism (Choi & Varian, 2009), or even influenza epidemics (Ginsberg et al., 2009).

Using the amount of search queries over the Internet for measuring the investor attention has advantages, such as the daily frequency and the fact that they reflect the real interest of the investor on the information (different from passive measures such as advertising expenses or published news). Besides, performing a search over the Internet is an action accomplished spontaneously. Therefore it is not a direct result of financial market activities, which limits endogeneity problems in the models.

The main challenge on this investigation is in the difficulty in measuring the attention levels of different types of investors to different types of information. With respect to searches performed over the Internet as active measures of attention, less sophisticated agents tend to gather information using popular and free-of-cost search engines, specially Google. Figure 1.1 evidences how most of the searches performed by Internet users in Brazil are by far performed using Google.

On the other side, professional investors tend to use paid-for platforms, specialized and with additional features, such as Bloomberg. Selecting the most adequate measure allows verifying whether the effect of attention in prices depends on who is attentive and to which kind of information.

Recent studies began to use Google search queries as an attention indicator of non-professional investors, finding interesting results. Kristoufek (2013) tested a diversified strategy that penalizes stocks with high search levels. The performance of this strategy was higher than the one of market indices. Preis et al. (2013) also found evidence that portfolios built according to search intensity for specific keywords present superior returns.

In this context, we aim to address the problem of the limitations of the attention as a cognitive resource and their implications on the Brazilian stock market. As far as we have searched, there is still a lot to be investigated about the influence of investor attention measured by search queries. This gap is even more evident in Brazil, although it has the 12th largest capital market in the world (Ministério do Planejamento, 2015).

This document combines the importance of attention in the capital markets with the increasing availability of Internet data that allows meaningful insights. Some of them are excellent proxies of the attention levels of investors. Hence, we developed three essays regarding the influence of attention fluctuations in the Brazilian stock market.

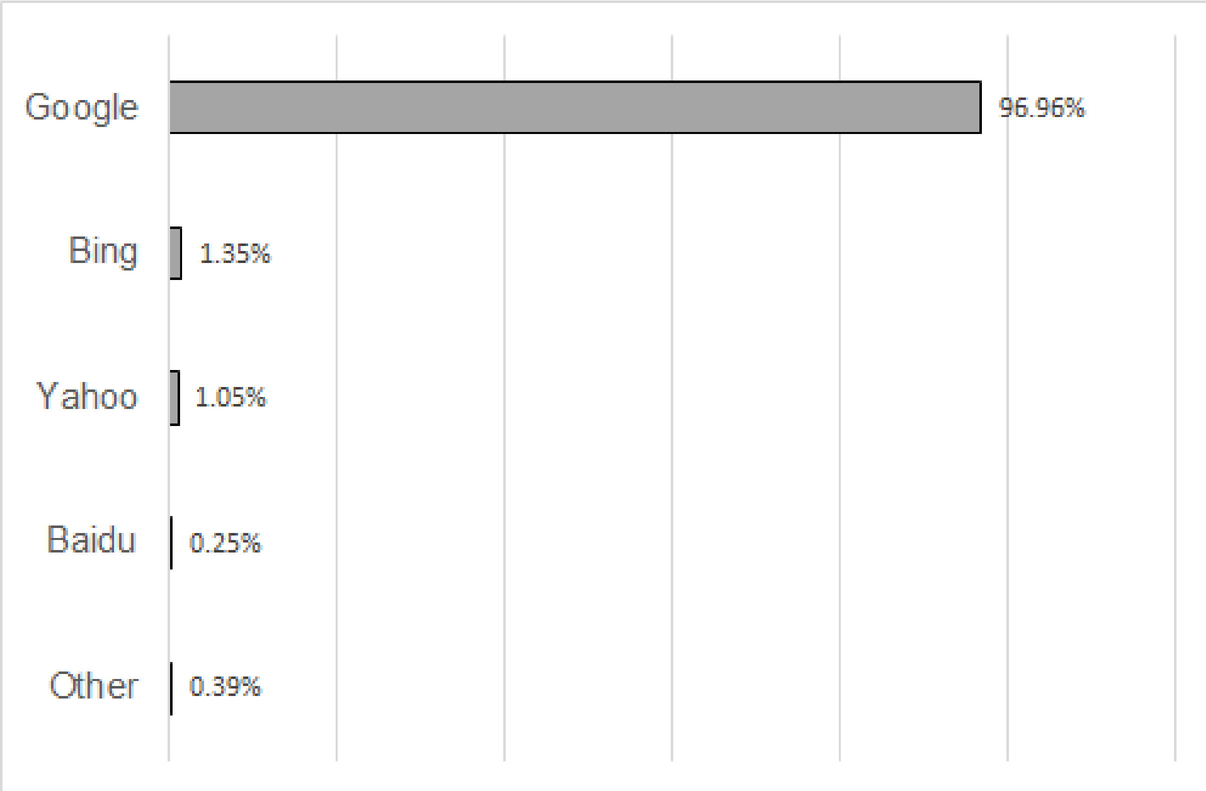

Figure 1.1: Search engines market share in Brazil
Note: Adapted from Statcounter. Period from Jan 2009 to July 2019.

Firstly, we verified whether markets are more efficient when investors are more attentive to them. We explored the possibility that deviations are more predictable when there is not enough attention to incorporate all available information into prices. After that, we assessed how attention relates to the volatility levels of the market, making a parallel between theories about noise trading and buying pressure. Our last essay approaches the relationship between attention and volatility asymmetry. It is well known that volatility is higher in bad times, but there is much to discover about the influence of attention fluctuations in this stylized fact.

All essays are structured with an introduction that outlines the research idea and their contributions, a revision of more relevant previous literature regarding this idea, and a description of the sample and the models developed for the hypotheses tests. Following that, we present the empirical results of the implementation of the models and we conclude each chapter with final comments.

These essays permit a better understanding of the practical effects of the attributes and properties of the investor attention. The approaches are in line with growing research with respect to the impact on markets of psychological aspects that challenge classical theories. A higher comprehension of the market dynamics contribute to investors improving their conditions to balance their portfolio according to their preferences, and to corporate officers managing public offerings and the release of information in conferences

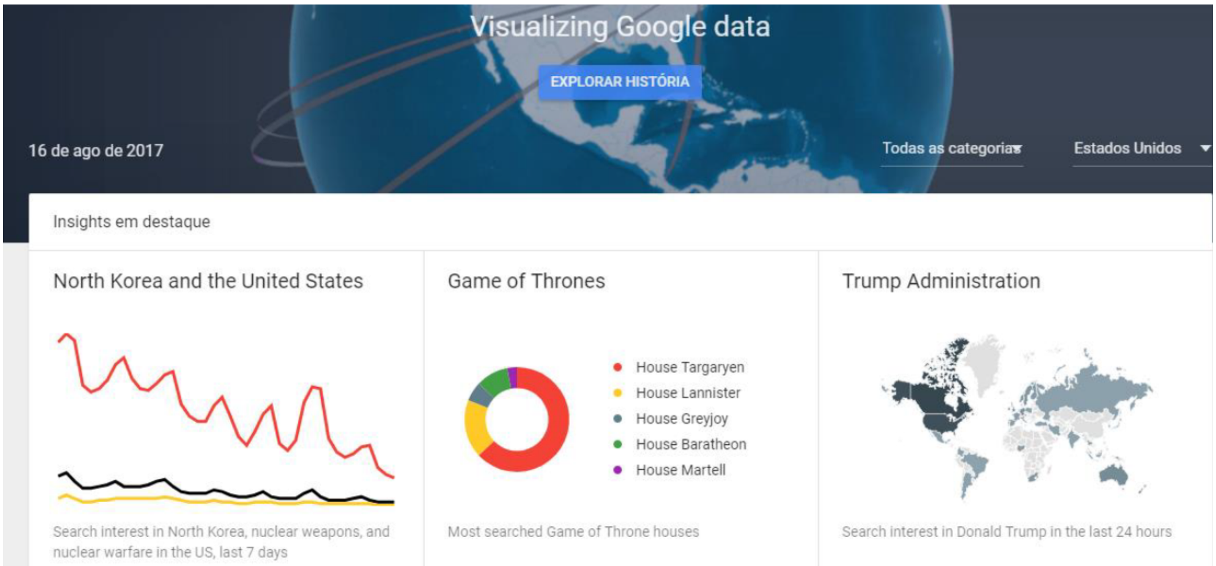

Figure 1.2: Google Trends
Note: Google Trends main page in Brazil: trends.google.com.

and institutional documents.

This study is also justified because it verifies the ability of recent measures that emerged from big data to describe or anticipate market behavior. Although developing tools and algorithms is not in the scope of this work, verifying the effectiveness of those measures subsidizes activities of trading, securities analysis and modeling, risk rating, portfolio management, among others.

In the next section we summarize the measures we used along the three essays to investigate the effects of fluctuations in investor attention.

1.1 Attention Variables

Three variables are used in this study to represent the attention of the investor. The first one is a proxy of the attention of the non-professional investor: an index of the amount of search queries performed at Google about financial assets. The second one is based on searches that users do at Bloomberg terminals, representing the attention of professional investors.

The last indicator represents the non-professional attention after filtering out the variance explained by professional attention, in order to better investigate the sole effect of retail attention in market dynamics.

1.1.1 Retail Investor Attention: Google Search Volume (GSV_t)

Google Trends provides search volume data since January 2004. Figure 1.2 shows the Google Trends website with the field where queries are executed.

In each query, the user can extract five keywords at most. Equation 1.1 describes how the tool calculates the index that is representative of Google search volume (GSV):

$$GSV_{b,r,t} = \frac{SVT_{b,r,t}}{TSV_t \times MSV_{r,\tau}} = \frac{SVT_{b,r,t}}{TSV_t \times max_{r,\tau}(\frac{SVT_{k,r,\tau}}{TSV_\tau})} \tag{1.1}$$

The indices b, r e t represent search term, geographical region where the search was done and period, respectively. The variable $SVT_{i,r,t}$ is the total search volume of the keyword i originated in the place r over the period t. TSV_t represents the total Google search volume over the period t and $MSV_{r,\tau}$ is the maximum ratio of searches for all keywords k included in the Google Trends query over the period τ. This calculation allows the indicator to eliminate temporal trends on the general use of the search tool. In other words, it eliminates the effect of an increase in the search volume of a specific keyword due to an increase in the number of Internet users. Besides, the volume is normalized to a range from 0 to 100 so to the point in time (month, week or day, depending on the query interval) with higher search volume will be assigned a GSV of 100. This method hampers comparisons between GSV series obtained in different queries (Welagedara et al., 2017). We considered only searches made in Brazil, to avoid effects related to timezones.

1.1.2 Professional Investor Attention: Bloomberg Search Volume (BSV_t)

While retail investors rely on free search engines, investment analysts and more sophisticated traders access financial data providers to find historical data and forecasts, as well as real-time quotes. As a critical part of the workflow of financial professionals, those one-stop-shop platforms offer up-to-date and accurate data.

Surveys indicate that Bloomberg has the highest share among platforms for finance research in the world. Its market share was 33% in 2018, equivalent to more than 300 thousand terminals currently licensed, more than competitors such as Capital IQ, Thomson Reuters Eikon and Factset (Institute, 2019; Times, 2018). Important features of the data provider are the instant messaging service and fixed income sections. The annual cost of a Bloomberg terminal is about 20 thousand dollars, although there are significant incentives to universities and labs. Figure 1.3 shows the share of Bloomberg terminal users by job title and a typical screen displayed by the terminal.

Breaking down Bloomberg users by job title evidences that a major part of them is institutional or professional investors. Portfolio managers, buy-side analysts and sell-side finance professionals are among the predominant Bloomberg users.

Although we did not find similar surveys locally focused in Brazil, we do not have reasons to think that Bloomberg is not one of the providers that dominate the financial

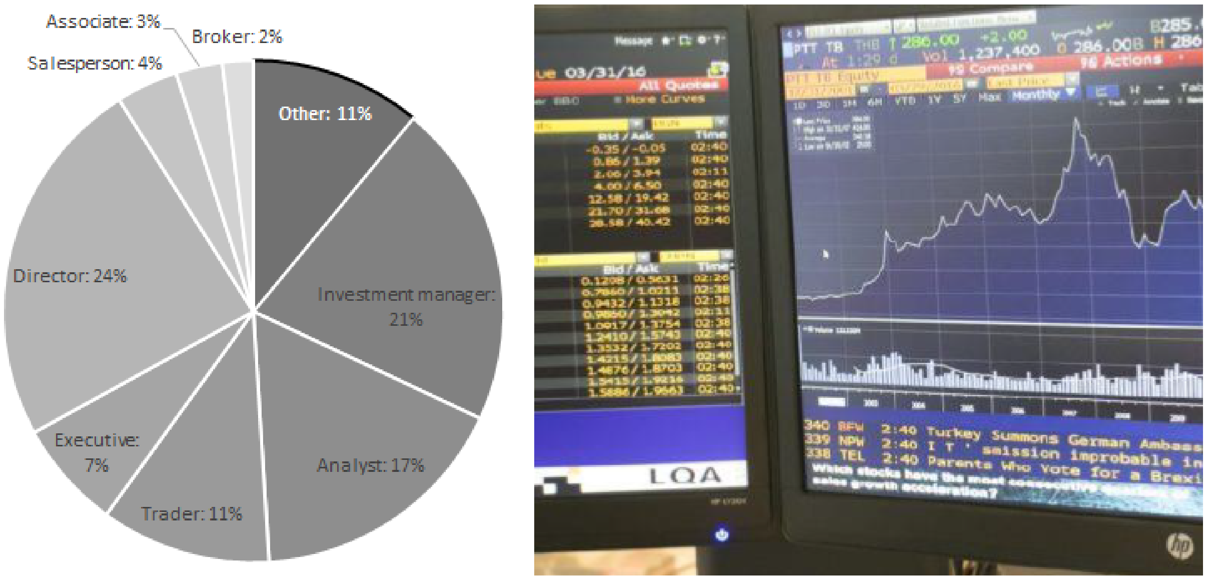

Figure 1.3: Bloomberg share of users by job title and Bloomberg terminal
Note: Adapted from Ben-Rephael et al. (2017) and www.coindesk.com.

data industry. We do know that some local competitors, such as Economática, may also have a relevant share.

Therefore, we obtained data on institutional attention on the Bloomberg terminal, using the same method adopted and explained by Ben-Rephael et al. (2017). We used the Bloomberg variable News Heat - Daily Max Readership, available at daily frequency. It is the maximum number from the hourly measures that are available in real time only.

The News Heat represents the number of times each article is read by its users and the number of times users search for a specific stock. There is no measure of News Heat for market indices. While searching requires users to type the stock ticker, reading does not imply specific interest in that firm. Hence, Bloomberg assigns a score of 10 for news searching and 1 for reading.

The variable ranges from 0 to 4, based on ranks comparing rolling averages to the user behavior during the previous 30 days. The rolling average considers the hourly counts during the previous 8 hours. Bloomberg assigns 0, 1, 2, 3 and 4 if the rolling average is (i) in the lowest 80%, (ii) between 80% and 90%, (iii) between 90% and 94%, and (iv) in the highest 4% of the hourly counts over the previous 30 days, respectively. After that, Bloomberg takes the maximum of all hourly scores in each trading day.

1.1.3 Isolated Non-Professional Attention: Google Search Volume Residuals ($rGSV_t$)

We adopted an additional attention measure to complement the analysis based on searches performed using Google and Bloomberg. It was considered because there is naturally a correlation between the attention of professional and of non-professional investors.

We call this measure the Isolated Non-Professional Attention ($rGSV_t$). It consists of the Google search variable after removing from its variance the variance explained by searches at Bloomberg.

Hence, $rGSV_t$ are the residuals of the regression of non-professional attention on professional investor attention, as Equation 1.2 shows.

$$\begin{aligned} GSV_t &= c + \delta BSV_t + \eta_t \\ rGSV_t &= \eta_t \end{aligned} \tag{1.2}$$

Using $rGSV_t$ allows evaluating more reliably the effects on markets dynamics arising precisely from the attention of the retail investor.

For each essay, the following chapters describe specific implementation decisions regarding the attention variables, the other variables of interest and the methods we apply. We also report the literature based on which the research hypotheses are developed and interpret the results.

References

Andreassen, P. B., & Kraus, S. J. (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting*, *9*(4), 347–372 (cit. on p. 10).

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *The Review of Financial Studies*, *21*(2), 785–818 (cit. on pp. 10, 20, 25, 28, 43, 47, 50, 51, 68, 86, 98).

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009–3047 (cit. on pp. 10, 15, 43, 88, 97).

Burton, E., & Shah, S. (2013). *Behavioral finance: Understanding the social, cognitive, and economic debates*. John Wiley & Sons. (Cit. on pp. 10, 48).

Choi, H., & Varian, H. (2009). Predicting initial claims for unemployment benefits. *Google Inc*, 1–5 (cit. on p. 11).

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499 (cit. on pp. 10, 11, 20, 43, 47, 48, 52, 53, 56, 57, 60, 65, 66, 68, 75, 87, 97).

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703–738 (cit. on p. 10).

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, *22*(2), 171–192 (cit. on pp. 10, 47, 48, 54, 58, 59, 65, 66, 68, 75, 86–88, 97, 98).

Gillings, M. R., Hilbert, M., & Kemp, D. J. (2016). Information in the biosphere: Biological and digital worlds. *Trends in ecology & evolution*, *31*(3), 180–189 (cit. on p. 9).

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012–1014 (cit. on p. 11).

Gomes, P., & Taamouti, A. (2016). In search of the determinants of european asset market comovements. *International Review of Economics & Finance*, *44*, 103–117 (cit. on p. 10).

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, *56*(4), 1533–1597 (cit. on pp. 9, 48, 49).

Huddart, S., Lang, M., & Yetman, M. H. (2009). Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science*, *55*(1), 16–31 (cit. on pp. 10, 47, 51).

Institute, B.-T. (2019). Financial market data and analysis global share & segment sizing (cit. on p. 14).

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ. (Cit. on pp. 9, 68).

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The Journal of Finance*, *63*(1), 273–310 (cit. on p. 10).

Kristoufek, L. (2013). Can google trends search queries contribute to risk diversification? *Scientific Reports*, *3* (cit. on p. 11).

Ministério do Planejamento, D. e. G. (2015). Mercado de capital brasileiro está em 12º no ranking mundial (cit. on pp. 11, 30).

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using google trends. *Scientific Reports*, *3*, srep01684 (cit. on p. 11).

Times, F. (2018). Bloomberg and reuters lose data share to smaller rivals (cit. on p. 14).

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141–162). Springer. (Cit. on p. 10).

Welagedara, V., Deb, S. S., & Singh, H. (2017). Investor attention, analyst recommendation revisions, and stock prices. *Pacific-Basin Finance Journal*, *45*, 211–223 (cit. on p. 14).

2 Information Discovery: Effects of Investor Attention on the Market Efficiency

No man is lonely eating spaghetti; it requires so much attention.

Christopher Morley (1890-1957)
Journalist, novelist and poet

The theoretical justification of the hypothesis of immediate incorporation of relevant information into asset prices is, to some extent, convincing. It is argued that incentives, information aggregation among investors and the activities of arbitrators eliminate the effects of limited attention.

However, some evidences have shown the opposite. Return autocorrelation or predictability and reversal or momentum effects are some examples. Alternative theoretical models have tried to explain those anomalies, such as Barberis et al. (1998) and Amihud (2002).

It is reasonable to imagine that these anomalies may also come from fluctuations in the absorption of information. Attention is a critical and time-varying resource. Psychological research documents that people can digest only a subset of information available due to time constraints, and this can lead to over- or underreaction.

Recently, some economic models tried to consider the idea that most people can easily find much more relevant information for making decisions than they in fact use to decide (Sims, 2006). This so called irrational attention is based on the assumption that agents have to select the information that is relevant among what is available. Grossman and Stiglitz (1980) are Merton (1987) are seminal studies that demonstrate the importance of the attention on asset price dynamics.

Several papers have addressed this question in recent years. The model of Falkinger (2008) determines that companies compete firstly for the attention of individuals, and then they can compete for their propensity to consume. This framework contributes to the explanation of already documented patterns, such as the home bias (Mondria & Wu, 2010; Van Nieuwerburgh & Veldkamp, 2007) and the equity premium puzzle (Gabaix & Laibson, 2005).

In this context, we investigated how our limited capacity of processing information affects stock prices in this study. If the level of attention to so much information is decisive for the behavior of financial markets, it is essential to examine whether an effect is transmitted to the predictability of these markets.

Excess of attention coming from non-professional investors may create extra noise and volatility (Barber & Odean, 2008; Da et al., 2011). On the other hand, more attention might mean more information transmitted to the market, increasing its efficiency. Our analysis departs from this last conjecture, named information discovery hypothesis (Fang & Peress, 2009; Grullon et al., 2004). It determines that more information absorbed and reflected in prices makes markets essentially less predictable.

Our study contributes to the behavioral finance literature, being a pioneer in analyzing the effect that the Internet activity induces in Brazilian stock prices, in a daily frequency. Besides, we did not find previous studies that compare the effect of access to relevant financial information providers on the market efficiency. The results allow a better understanding of how the increasing abundance of data and democratization of the technology can make markets more predictable. If this effect exists, entities whose mission includes promoting the development of the stock markets should stimulate higher usage of those information providers, as well as advocate for better disclosure by companies and other market agents.

We aimed in this study to answer whether the attention of professional and non-professional investors can affect the predictability of the Brazilian stock market. Our findings may help companies that want to promote their recognition by investors. Understanding how attention influences the incorporation of information into prices contributes to a better management of the launch of releases and ads. Moreover, this effect can define the visibility desired by the firm, particularly on the Internet. Furthermore, other market participants (eg., arbitrageurs) might benefit from market anomalies measured by the attention level of a specific class of investors.

Our study is structured as follows: in the next section, we present the underlying theoretical background, as well as related previous analysis. After that, we describe the methodology implemented, analyze the results and present final considerations.

2.1 Literature Review

2.1.1 Information Discovery Hypothesis

Commercial transactions occur between individuals due to heterogeneities of preferences, wealth or beliefs. Focusing on differences in beliefs, Grossman and Stiglitz (1980) presented some conjectures describing the nature of equilibrium when prices transmit information. They claim that a competitive market cannot be always in equilibrium, because that would take out any possibility for arbitrageurs to profit for their costly activities. Therefore, they conjecture that prices only partially reflect the information of arbitrageurs. This makes gains feasible, creating the possibility for compensating the expenses for obtaining information with those gains. Besides, the information level of the markets depends on the amount of informed individuals.

The market is, therefore, a system in which prices, through buy and sell orders, exert a role of transferring information from the more to the less informed. However, there would not be any equilibria if this process occurred perfectly.

The model of rational expectations with noise of Grossman and Stiglitz (1980) considers two assets: a riskless one and a risky one. The return of the risky asset has two components: the first one can be observed at a cost and the second one cannot be observed. There are also two types of agents: the informed ones, which observe the first component of the return, and the ones that see only the price. *Ex ante*, all the investors are identical. Whether they spend to obtain information is what differentiates them.

Informed agents will trade depending on the price and the observable component of the return. Non-informed ones trade only according to the price, but have rational expectations. The price is, hence, a function of the non-informed component of the return and of the inventory of the risky asset. The equilibrium quantity of informed traders is a function of this price and the cost of obtaining information.

Uninformed traders do not know the amount of this inventory neither are able to completely differentiate the observable component of the return by watching the prices, because their deviations might occur due to fluctuations either in the information absorbed by the informed ones or in this inventory.

Traders convert from one class to the other depending on the expected utility of each class (considering the cost for obtaining information). In general equilibrium, the utility of these two classes converges, since the expected utility (or the gains) of the class of informed ones decreases with the increase of the number of informed investors. This occurs because when more traders see the observable components of the return, fluctuations in this return induce a larger effect on the aggregate demand and, consequently, on the price. With this, a larger part of the information obtained by informed ones becomes available for non-informed ones, increasing the informativeness of this price system. The gains for informed ones are proportional to the difference between market prices and its value in case all information were equally distributed. In general terms, the gain that informed investors obtain from non-informed ones is distributed to the amount of informed traders.

The quantity of informed investors increases the informativeness of the price system, decreasing the volume of noise that distorts the information transmitted through this system. In the hypothetical situation of the absence of noise, prices would transmit all the information, thereby not having incentives to obtain information. Markets would be scarce. Hence, there would be an equilibrium with the absence of information. However, if nobody is informed, there would be a reward for the acquisition of information, preventing a competitive equilibrium. An equilibrium would require identical beliefs among all when everybody (or nobody) is informed.

The seminal work of Grossman and Stiglitz (1980) is base for a paradigm that challenges the idea that prices immediately reflect the entire available information. In a process of incorporation of information based on fluctuations in the amount of informed and in their expected utility, the attention level of the agents might fulfill an essential role.

Merton (1987) also criticized assumptions of the classical theory. The author questioned traditional premises, such as the one that companies are capable of raising immediate and enough capital for investments and the small importance given to financial intermediaries.

The process of acquisition and distribution of information is profoundly analyzed. Pricing models generally assume that the diffusion of public information is instantaneous for all the investors and raise immediate initiatives. However, in practical terms, that depends heavily on the nature of the information and the time scale of the analysis. It is reasonable to imagine that information released by standardized channels, such as performance announcements, will generate the timely reactions expected in these models. The same cannot be told about a profitable opportunity due to an empirical anomaly discovered and presented in a scientific journal.

Depending on the scope of the anomaly, considerable time and amount of investors may be necessary to correct it. After staying aware of it, an investor will firstly ask himself whether it would maintain in the future and whether potential gains would be enough to compensate the implementation costs of the strategy. In the same way, authors of these papers shall verify whether, during the whole period used for statistically identifying the anomaly, the suitable technologies to analyze the data were available.

In this context, Merton (1987) proposed a market equilibrium model with heterogeneity in the absorption of information by the investors. The model considers that an investor is informed about an asset if he knows its parameters (expected return, physical investment rate and production technology constant). The basic condition for considering this asset in his portfolio allocation is knowing about it.

As Grossman and Stiglitz (1980), Merton (1987) presents a model that conjectures that all the informed investors agree with the values of the asset parameters. However, the fact that only equally informed investors trade a specific asset implies no information asymmetry. Moreover, this model considers that the information quality about assets is the same for any asset, focusing on the unequal distribution of this information among investors.

Merton (1987) segregates the costs of releasing information (incurred by the companies) in two components: the cost of seeking and processing data and the cost of transmitting it. The first one is marginally low since these data are already necessary for managing the company. However, agency and signaling theories evidenced that the cost of transmitting information might be considerable for being efficiently used by investors. These costs comprise the necessary incentives for managers to transmit information and the efforts to make it reliable.

In addition to these costs, it is necessary that the investors incur in expenses even before processing the information made available. In order to receive the information it is necessary that they follow that firm in particular. Since for current shareholders this fixed cost is sunk, the information received by them is different than that one received by potential investors. This logic also holds for information released by other sources, such as portfolio managers and analysts.

The proposed model supports the concept of neglected stocks, the ones that are not followed by a significant number of analysts. Consequently, less information about those firms is made available, resulting in lower equilibrium expected returns of their stocks, compared to other ones.

Merton (1987) started from the premise that all investors have identical preferences and same initial wealth. Considering the market portfolio as the weighted average of the optimal portfolios of the investors and that all investors adopt the same exposure to the common factor, the exposure to the market portfolio to this factor will be the same. In equilibrium, the market value of an given firm will always be lower with incomplete information, being the effect similar to an additional discount rate. The larger the investors base, the higher will be this difference. The additional expected return is proportional to the hidden cost of this incomplete information. The market portfolio is not efficient with respect to mean-variance in this incomplete information model.

Assessing the effect of the investment base on the equilibrium expected return of a stock is more difficult for larger or more known firms. One of the reasons is that their total variance and its fraction associated with idiosyncratic volatility are much lower for these type of stocks. More relevant is the fact that a higher percentage of the stocks of those firms is held by financial institutions (such as investment funds), resulting in an underestimation of the effective amount of individual investors.

In the case of stocks of relatively small firms and with a smaller base of institutional investors, the model with reported data is capable of estimating a expected return differential, in equilibrium, compared to the corresponding value of the model with complete information.

More recent studies based on these seminal papers examined how the attention level contributes to the process of absorption of information. Sims (2006) highlighted the inertia in the reaction of economic agents to the release of external information due to the limited attention capacity. This inertia is typically represented in economic models by adjustment costs, or by releasing and implementing delays. The literature that addresses these questions commonly presents limitations such as assuming that prices are observed with errors. This is equivalent to assuming unrestricted capacity of processing information.

Still today, one may question whether advances in identifying empirical anomalies will result in discarding the rational behavior paradigm, or whether that issue will be solved in this traditional framework. More recent papers adopted empirical approaches to analyze the effects of this gradual process of incorporation of information into prices.

Grullon et al. (2004) did an analysis in this sense, supposing that the impact of the firm visibility among its investors might have implications in the capital market. Based on this hypothesis, they verified whether the amount expended by the company with advertising affects its stock liquidity and the investors base. Their findings, robust to different methodologies, evidenced that more visible companies present stocks with better liquidity and have more individual and professional investors in its public float.

Fang and Peress (2009) verified whether there is a cross-sectional relationship between media coverage and stock returns. The authors obtained data from 1993 and 2002 of stocks listed in New York Stock Exchange (NYSE), mostly of large firms. Media coverage was measured by counting the number of articles in influential newspapers of mass circulation. They represent around 6 million daily printed copies, 11% of the American market.

Only relevant articles were considered, according to criteria related to frequency, representativeness and the position of keywords. The data showed that coverage is on average low and asymmetric, besides being homogeneous among segments. The coverage depth varies significantly among the considered newspapers, although there is a relevant overlap. Both stocks with and without coverage tend to remain in the same situation in subsequent periods.

Applying Fama-MacBeth regressions, the authors found that the firm size is determinant of media coverage, as expected. Besides, companies with low book-to-market ratio and with more analyst coverage tend to have less media coverage. Firms with more individual shareholders, opinion dispersion among analysts and idiosyncratic volatility tend to have more coverage.

The univariate analysis evidenced a strong and negative relationship between media coverage and returns. Dividing stocks in tertiles according to the coverage level and equal weights, the authors found an annual differential return of 4.8% between stocks with less and stocks with more coverage. This lack of coverage premium remains after controlling by several characteristics, such as size.

The authors then regressed the returns of long & short portfolios on the Fama-French, Carhart and Pastor-Stambaugh factors. The resulting alphas confirmed the lack of coverage premium. It is partially explained by the factors, however. This long & short strategy presented positive correlation with the stock returns of small, value and high momentum companies, and negative correlation with the market returns. This pattern persisted in the analysis that considered the characteristic-based benchmarks of Daniel et al. (1997). The results indicated that the effect is predominantly determined by the long position in stocks with less coverage. This finding challenges the conclusions of Barber and Odean (2008), that this effect is the result of a buy pressure of stocks that call more attention from individual investors.

Dividing the sample in subsamples according to size, book-to-market and returns over the last 12 months, the results showed that the effect is more expressive among stocks of companies that are smaller, and that have lower book-to-market ratios and previous returns. Excluding stocks with very low prices or trading volumes, a robustness test, indicated that the explanation cannot be related to microstructure problems, such as the bid-ask bounce.

Additional tests also rejected hypotheses related to drifts after performance announcements, as well as to post-Initial Public Offering (IPO) or tech companies underperformance, when those cases were excluded from the sample.

Fang and Peress (2009) then verified three factors that might explain the media coverage effect found: continuation or reversion patterns of returns, trade prevention, and investor discovery. The first one, documented by Chan (2003), is associated with the evidence that stocks with low returns and evident in news show negative drifts in the following 12 months. On the other hand, the ones with low returns but no visibility in the news show short-term reversions subsequently. This alternative explanation was discarded because the long (and not the short) component of the long & short strategy is the determinant of the media exposure effect found. If this effect were defined by the short element, a negative drift common for high-exposure losers could be assessed as a reason. The authors also discarded the possibility that the effect could come from short-term reversions registered by low-exposure losers when they verified that the alphas of this long & short strategy persisted for considerably more than a month.

The hypothesis of trade prevention is based on the conjecture that this media exposure effect cannot be adjusted due to illiquidity. After grouping the stocks by different liquidity measures, the results were controversial: the effect was larger for the less liquid stocks when measured by price and buy-sell spread, but this pattern did not occur after measuring by the illiquidity ratio of Amihud (2002) or by daily trading volume. Thus, the authors concluded that illiquidity shall contribute to the persistence of the exposure effect, but not to its emergence.

Informationally incomplete markets require that stocks recognized by few investors compensate them for the under-diversification. Fang and Peress (2009) grouped the stocks by quantity of analysts and individual investors, since these measures indicate the level of dissemination of the information about the firm. The same is done with the idiosyncratic volatility, since it should be compensated by higher returns due to the undiversified risk it imposes. The results indicate that the exposure effect is more pronounced for companies with less analyst coverage, higher concentration of individual investors and higher idiosyncratic volatility.

The information discovery hypothesis is based on a non-immediate process of incorporation of news into asset prices. One of the main reasons for this particularity is the scarcity of the attention as a cognitive resource of the agents. Some studies analyzed the effects of this condition.

2.1.2 Competition for Attention

Falkinger (2008) developed a framework for the interactions between the attention scarcity and the conventional economic scarcity. In this model, the competitive equilibrium and important macroeconomic variables are modeled based on empirical facts about attention and psychology of perception.

The central point of the model of Falkinger (2008) is that psychological parameters and factors both economical and of information technology determine whether an economy is rich or poor in information volume. In the first case, only economic properties define profitability, thereby there are no obstacles for a specific item to reach consumer's attention. On the other side, an economy that is rich in the amount of information is the one in which technologies, globalization and distribution over media channels increase the level of attention-demanding activities and the sign force that is necessary to capture the attention of recipients.

The basic assumption is that the economic modeling should consider psychological facts. According to the author, psychological studies identified two aspects of the human information processing: a selection of a set of items to be processed based on a filter mechanism and the processing of the selected items through a limited allocation of cognitive resources. This filter mechanism implies that the cognitive resources for processing are allocated only on the items that passed through this gating, while the other ones are ignored.

Some aspects of the model address the allocation of cognitive capacity. The speed and quality of the information processing depend on this available capacity after the load related to the exposure to other signals. This capacity relies on the effort of the recipient, and increases when a higher exposure to a signal increases the level of arousal. The total exposure is the sum of the signals sent by the attention-seeker agents. The attention attracted by the force of the agent's signal determines its cognitive impact in the recipient. Events that capture attention have more tendency to be consciously perceived, and in detail. Besides, they have more chances to elicit controlling responses and to be stored in the permanent memory.

The fundamental constraint pointed in the model is that any behavior, regardless of the rationality level, is contingent on a perceptual filter. This filter determines which agents will have their signals captured by the mind of the recipient. It imposes restrictions in the individual behavior and in the market interactions, which complement the traditional restrictions related to budget and resource. Two types of triggers are involved in this process: the first one is related to the absolute force of the signal, even in an environment totally free of information sources. There is also a relative trigger, in which a source should be noticeable with respect to others that compete with other signals.

The model of Falkinger (2008) also describes the behavior of this perceptual filter when there is intermediation through information channels, such as search engines or mass media. In these cases, the attention given depends on the salience level of the news. In other words, it relies on the emphasis given to it in the media. For instance, readers typically see a limited number of results provided by a search engine. Regardless of psychological gate, the media gate requires that the issued signal have a minimal force to reach the mind of a desirable amount of recipients.

The psychological impacts of the buying behavior stemming from a information-rich economy result in superior margins or in a bias in favoring more attention-competitive sectors. The wealth is higher in this type of economy. Individuals become more oriented to consuming and to labor when changes in economic fundamentals raise the degree of informational wealth of the economy, intensifying the competition for scarce attention. On the other hand, if the factors that result in an elevation in the competition for attention do not increase labor productivity, wages and income deteriorate. In an economy with low volume of information, there is an decentralized competition for attention and wealth results in an efficient equilibrium, whereas in an information-rich economy the resulting equilibrium is inefficient due to an exacerbated competition for scarce attention.

Attention levels impact the economy and, consequently, the financial markets. The activity level in the Internet is an important measure of investor attention since it is one of the major means of actively obtaining financial information. In the next section, we analyze the main studies that, based on this proxy, addressed the effects of the attention on the predictability level of the prices of publicly traded assets.

2.1.3 Attention and Market Efficieny

One of the first papers to investigate, empirically and using Internet data, the influence of attention on market efficiency was Vozlyublennaia (2014). In an initial approach, the author verified that the attention given to an index has a significant short-term effect in its return. This is consistent with the hypothesis that uncovered information can be perceived by investors as indicating a higher (or lower) future return, resulting in an increase (or decrease) in the security returns. This evidence is different from the idea that an increase in the attention would lead to a buy pressure and therefore to price increases, established by Barber and Odean (2008). The opposite effect (effects of return shocks on attention) was evidenced in the long term.

She examined the dynamics between investor attention and the performance (returns and volatility) of six different asset indices: stocks (Dow Jones Industrial Average, S&P 500 and National Association of Securities Dealers Automated Quotations (Nasdaq)), bonds (Chicago Board Options Exchange 10 Year Treasure Note Yield Index), commodities (West Texas Intermediate Crude Oil) and gold (Chicago Board Options Exchange Gold). The attention was measured by the search volume of each index at Google. The analysis comprised weekly data from 2004 to 2012.

The author measured the effect of the attention on the predictability of returns by including interaction terms of past returns and attention. The results evidenced significant effects in these interactions and, more importantly, opposite to the autocorrelation effect. This means that the attention reduces return autocorrelation. In other words, attention shocks reduce return predictability, consequently increasing market efficiency. The regressions were controlled by possible changes in investment opportunities using macroeconomic variables.

The results indicated that the interaction terms lead to a reduction in the predictability of the returns of Dow Jones Industrial Average (Dow) and S&P 500 indices, more representative of large and medium companies. This pattern was also verified in the crude oil index. The relationship was not significant for bond and gold indices. In the Nasdaq index, however, the same sign was found in the autocorrelations and the interactions.

As well as for returns, similar analysis were developed for volatilities. Shocks in attention lead to a decrease in the predictability of the short-term volatility in Dow, Nasdaq, gold and crude oil indices.

In an additional analysis of cross-correlations, the author verified that the attention decreases the capacity of the stock returns of large companies to predict the returns of small ones, as well as of gold and oil. On the other side, attention reduces the predictability of returns of larger companies (representative of the Dow index) through the returns of bonds and gold indices.

Tantaopas et al. (2016) also investigated the relationship between attention and market efficiency examining the impact that the amount of search queries over the Internet might have in market variables (return, volatility and trading volume). The authors addressed developed and developing markets in Asia-Pacific. The literature about such effects in these countries is scarce. They obtained data of Australia, Hong Kong, Japan, New Zealand, Singapore, South Korea, China, India, Malaysia and Thailand. These countries account for almost 28% of the global GDP.

The information discovery hypothesis, according to the authors, follows a recurrent process that starts with an abnormal return due to a market reaction to some corporate event (such as earnings releases). If this event is sufficient for capturing the attention of some investors, they will seek information about the company stock. In case they find good (bad) news, they will make the decision to buy (sell) that stock, which might result in a price increase (decrease). According to them, the positive price pressure due to abnormal attention does not explain price decreases and does not hold when short selling is allowed.

The authors obtained weekly data of search volume at Google from 2004 to 2014. They extracted the indices that comprise a representative share of the total market, one from each country. The market efficiency was measured by the predictability of the returns, volatilities and trading volume. As well as Vozlyublennaia (2014), they adopted lagged interactions between these series and the search volume.

A specification using autoregressive vectors indicated a decrease in return predictability in six countries: Hong Kong, India, Malaysia, Japan, Korea and Singapore. In these latter three countries, increasing attention contributed to a predictability reduction of 34.1%, 35.8% e 36.8%, respectively.

With respect to volatility, Internet searches contributed to a reduction in the predictability in eight among ten countries (the pattern was not evidenced only in New Zealand and Thailand). In China, the decrease in predictability was by 56.1%. In general, the reduction in return predictability in developed countries was stronger than in developing ones. The opposite was found for volatility predictability.

On the other hand, with the exception of Singapore, a significant effect was not found when testing the capacity of attention to reduce the predictability of trading volume.

The authors also concluded that the majority of the causality relationships are one-way: changes in return, volatility and trading volume determine changes in attention. Whether the relationship is direct or inverse depends on how much time has passed since the impulse, as the impact is short-lived for both developed and developing countries.

To a certain extent, these papers presented convergent results assessing the effect of attention levels in the efficiency of financial markets. Figure 2.1 presents the logic behind this pattern. Fundamental shocks capture the attention of investors, who will search (mostly on line) for more detailed information for trading decisions. Buying and

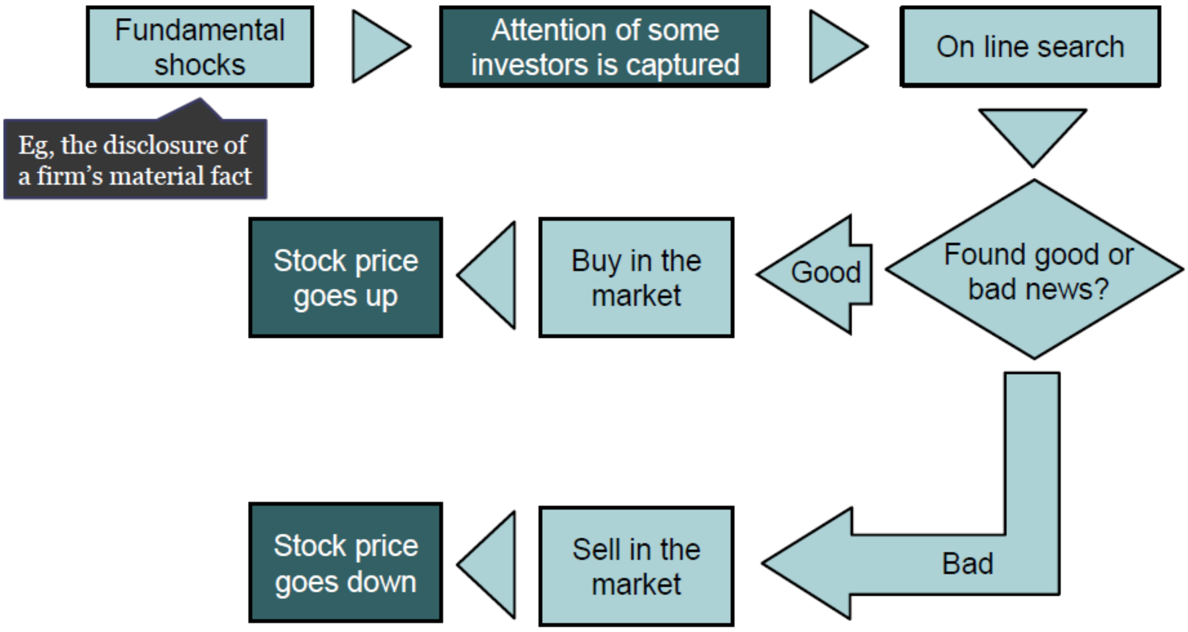

Figure 2.1: Flow diagram of the role of attention in the incorporation of information into prices

Note: Based on the theory of Tantaopas et al. (2016).

selling behavior will determine price movements. Despite somewhat trivial, this flow evidences that attention is required for the incorporation of fundamental news on prices through trading. Hence, the more attentive the market is, the more immediate might be the market movements, resulting in less autocorrelation, less predictability and more efficiency.

This chapter innovates by investigating Brazilian individual stocks in a daily frequency. Besides, we verify two distinct classes of investors: professional and individual ones. This approach is useful to verify whether this pattern holds in this important developing market (Ministério do Planejamento, 2015). Hence, we start from the hypothesis that higher degrees of attention coming from some type of investors lead to lesser return predictability:

H_1: increased investor attention promotes market efficiency

The next sections address the specifications and the data with which the hypotheses analysis was performed.

2.2 Methodology

In this section, we describe how the hypothesis of attention-induced reduction of market predictability was tested. We provide details on sources and features of financial and search data, as well as the empirical model and expected relationships.

2.2.1 Data Description

We opt for analyzing the universe of individual stocks since, although the literature on the influence of the Internet activity in stock indices is scarce, it is even more scarce for stocks. Besides, the measure of professional investor attention adopted in this study is available only for this type of securities. This way, we chose to analyze the 67 stocks that comprised the Ibovespa during the second third of 2018.

The Bovespa Index (Ibovespa), the most popular Brazilian stock index, was created in 1968 and represents a reference for investors worldwide. The stocks that compose the index correspond to around 80% of the traded volume of the Brazilian capital market (B3, 2015).

The portfolio of stocks that will be part of Ibovespa is redefined every four months. The criteria for this process takes into account mainly the trading volume and attendance. The assets are weigthed by the market value to compose the index theoretical portfolio, with certain rules and limits (B3, 2015).

Our analysis comprise indicators of the attention from both professional and non-professional investors. To obtain an attention measure of non-professional investors, we obtained series of searches executed at Google. Google is by far the most popular search engine in Brazil and in the majority of countries (Statista, 2014). Since it is free and it is not a specific tool for this type of analysis, it is used by non-sophisticated investors.

Hence, between 5th-11th September 2018, we obtained at Google Trends, individually, series of the search volume index of all the 67 stocks of the sample. Google Trends is a free tool that allows extracting normalized weekly values of search volume of an specific keyword given the period and the geographical location. When there is no weekly information available, the tool presents monthly numbers (Adachi et al., 2017).

In order to extract the series of the index representative of the Google search volume of each individual stock, we adopted as keyword exactly the ticker symbol of each stock. This strategy avoids the influence of searches made with non-financial or non-trading purposes, such as for product consumption or job opportunities.

Our study pioneered in analyzing the influence of attention on the daily predictability of returns. However, Google Trends does not provide daily data for periods that are larger than a quarter. Hence, we had to obtain daily series of each quarter separately and then concatenate them using the method proposed by Johansson (2014) from weekly and monthly series over the whole period. Basically, daily volume series are multiplied by a factor equal to the ratio of the weekly volume to the daily volume of the first day of the week.

Moving on to professional attention, we extracted Bloomberg search volume series for each individual stock that comprised the Ibovespa during the second third of 2018, to perform the analysis. We exported the trading days data to Excel using the proper software add-in.

For the daily series of both types of attention measures, we considered only trading days. As expected, we found several days with information not available. Hence, due to lack of sufficient data, some of the 67 stocks had to be dropped.

Specifically, three tickers were discarded for not having enough available information during entire quarters: B3SA3, EGIE3 and SMLS3. Moreover, 11 stocks presented less than 100 available observations at Google Trends and then were also discarded in the analysis: BRAP, CPLE6, CVCB3, GOAU4, ITSA4, KLBN11, LAME4, SAPR11, SUZB3, TAEE11 and VVAR11.

Besides, our approach disconsiders days with search volume equal to zero, since that hampers the conversion to a logarithmic scale and can indicate unavailable information. Hence, the tests involving Google search volume were done based on 26 stocks, while those involving Bloomberg search volume contemplated 38 stocks. Table 2.1 presents the companies and their respective stocks, grouped by segment and indicating which ones were part of each analysis. We marked in boldface the stocks with the largest share in the index, which totaled more than 43%.

After obtaining attention indicators, we extracted daily nominal dividend-adjusted log-returns and trading volume of each individual stock that comprised this index during the second third of 2018. The extraction was made on August 16, 2018. These series were obtained in the software Economática, from 2011 to 2017, a long enough period for this analysis.

2.2.2 Modeling

For this test we adapted the method applied by Vozlyublennaia (2014) and Tantaopas et al. (2016). When the market is informationally efficient, there is no serial dependency in asset returns (Fama, 1965). Serially dependent returns are, at least partially, predictable. Hence, more market efficiency implies less return predictability. A regression model with interaction terms is useful to verify how attention levels alter this degree of predictability. The autoregression Equation 2.1 relates current and past returns, as well as interaction terms between past returns and attention. Search volume (SV) is the proxy that represents attention.

$$r_{t,i} = c_i + \sum_{j=1}^{p} [\beta_{j,i} r_{t-j,i} + \delta_{j,i}(r_{t-j,i} \times SV_{t-j,i})] + \epsilon_{t,i} \tag{2.1}$$

For each stock i, we regress its time series of log-returns ($r_{t,i}$) on their lagged terms ($r_{t-j,i}$) and the interaction of these terms with the search volume of the same day ($r_{t-j,i} \times SV_{t-j,i}$).

Table 2.1: Sample of stocks

Sector	**Firm**	**Ticker**	**% in Ibovespa**	**Enough *GSV* data?**	**Enough *BSV* data?**
Aerospace Manufacturing	Embraer	EMBR3	1.18	Yes	Yes
Air Transportation	Gol Linhas Aéreas	GOLL4	0.17		Yes
Banking	Banco do Brasil	BBAS3	3.12	Yes	Yes
	Banco Bradesco	BBDC3	1.41	Yes	Yes
		BBDC4	**7.73**	**Yes**	**Yes**
	Itaúsa	ITSA4	3.49		
	Itaú Unibanco	**ITUB4**	**10.44**	**Yes**	**Yes**
	Banco Santander Brasil	SANB11	0.98		Yes
Beverage	AmBev	**ABEV3**	**7.06**	**Yes**	**Yes**
Chemical	Braskem	BRKM5	0.82		Yes
Drugs	Hypera Pharma	HYPE3	0.88	Yes	Yes
	RaiaDrogasil	RADL3	1.01		
Education	Estácio Participações	ESTC3	0.69	Yes	Yes
	Kroton Educacional	KROT3	1.42	Yes	Yes
Energy	Cosan	CSAN3	0.42		Yes
	Grupo Ultra	UGPA3	2.12		
Engineering & Construction	Cyrela Brazil Realty	CYRE3	0.24		Yes
	MRV Engenharia	MRVE3	0.30		Yes
Financial Servies	Cielo	CIEL3	1.46	Yes	Yes
	Brasil Bolsa Balcão S/A (B3)	B3SA3	3.57		
Food Processing	BRF	BRFS3	1.36	Yes	Yes
	JBS	JBSS3	0.95	Yes	Yes
	Marfrig	MRFG3	0.22	Yes	Yes
Healthcare	Fleury	FLRY3	0.57		
	Natura	NATU3	0.40	Yes	Yes
	Qualicorp	QUAL3	0.48		
Holdings	Bradespar	BRAP4	0.58		
Insurance	BB Seguridade	BBSE3	1.34	Yes	
Machinery	Weg	WEGE3	0.95		
Metals & Mining	Cia Siderúrgica Nacional	CSNA3	0.38	Yes	Yes
	Gerdau	GGBR4	1.18	Yes	Yes
	Metalúrgica Gerdau	GOAU4	0.32		
	Usiminas	USIM5	0.39	Yes	Yes
	Vale	**VALE3**	**11.23**	**Yes**	**Yes**
Oil & Gas	Petrobras	PETR3	4.00	Yes	Yes
		PETR4	**6.73**	**Yes**	**Yes**
Others	Localiza Hertz	RENT3	0.97		
	Smiles Fidelidade	SMLS3	0.31		
Paper & Forest Products	Fibria Celulose	FIBR3	1.15	Yes	Yes
	Klabin	KLBN11	0.90		
	Suzano Papel e Celulose	SUZB3	1.39		
Real Estate Development	brMalls	BRML3	0.64		Yes
	Iguatemi	IGTA3	0.22		
	Multiplan	MULT3	0.41		
Recreation	CVC Brasil	CVCB3	0.57		
Retail	B2W Digital	BTOW3	0.33		
	Lojas Americanas	LAME4	0.92		
	Lojas Renner	LREN3	1.65		Yes
	Magazine Luiza	MGLU3	0.48	Yes	
	Via Varejo	VVAR11	0.27		
	GPA	PCAR4	0.87		Yes
Telecom. Services	Tim Brasil	TIMP3	0.90		Yes
	Telefônica Brasil	VIVT4	1.45		Yes
Transportation	Grupo CCR	CCRO3	0.94		Yes
	EcoRodovias	ECOR3	0.13		
	Rumo	RAIL3	1.19	Yes	
Utilities (Electric Energy)	Cemig	CMIG4	0.47	Yes	Yes
	CPFL Energia	CPFE3	0.09		Yes
	Copel	CPLE6	0.19		
	Eletrobras	ELET3	0.38	Yes	Yes
		ELET6	0.37		Yes
	EDP Brasil	ENBR3	0.29		
	Equatorial Energia	EQTL3	1,00		
	Taesa	TAEE11	0.28		
	Engie Brasil	EGIE3	0.54		
Utilities (Water and Sewer)	Sanepar	SAPR11	0.28		
	Sabesp	SBSP3	0.84		Yes

Note: 67 stocks that comprised the Ibovespa in the second third of 2018, base for our analysis. Some companies have more than one stock class, with different rights. The representativeness in the Ibovespa market value is with respect to May 7, 2018. The last 2 columns indicate whether enough information (at least 100 observations) about each ticker could be obtained.

The coefficients $\beta_{j,i}$ indicate the relationship between $r_{t,i}$ and its autoregressors. On the other hand, $\delta_{j,i}$ indicates how attention affects this relationship. We theorize that attention reduces return predictability. Hence, whenever $\beta_{j,i}$ shows a significant effect of $r_{t-l,i}$ on $r_{t,i}$, we expect that $\delta_{j,i}$ also shows a significant but opposite effect of $r_{t,i} \times SV_{t,i}$ on $r_{t,i}$. In other words, we expect that higher levels of investor attention reduce the explaining power of past returns on current ones.

The variable $\epsilon_{t,i}$ represents the shocks, i.i.d. All the variables are in the logarithmic form in our setting. The order p is the maximum lag of the regressors, defined according to information criteria and parsimony.

Our setting also includes a specification with the trading volume ($TV_{t,i}$) as a control variable. Hence, Model 1 considers returns and their interaction with search volume, and Model 2 includes trading volume as a third regressor. We do not expect that the trading volume will alter the relationship between attention and return predictability.

As explained, we will use Google Search Volume ($GSV_{t,i}$) and Bloomberg Search Volume ($BSV_{t,i}$) as attention variables ($SV_{t,i}$). As an additional test, we will use the residual of the regression of ($GSV_{t,i}$) on ($BSV_{t,i}$) to verify the sole effect of non-professional attention. This residual specifically represents non-professional attention not captured by professional attention. Chapter 1 describes other aspects of these three attention variables.

2.3 Empirical Results

In this section we describe the sample, also presenting correlations among the different search volume indicators. After that, we present the results of the empirical analysis of the effects of the investor attention on the stock market efficiency.

2.3.1 Descriptive Analysis

Table 3.2 presents general features of the variables involved. Discarding cases with zero search volume, the sample encompasses 127,201 observations of trading volume, returns and search volume, both at Google and Bloomberg.

The measures evidence a daily average volume of around 10 thousand transactions. Daily average returns are very close to 0, with a standard-deviation of 2.67%. The largest trading volume for a single stock in a day was approximately 236 thousand, whereas the highest average daily trading volume throughout the whole period among all stocks was 40 thousand. The return and Google search volume time series stand out for their higher variability.

Firegure 2.2 presents the correlations among the search volume indices of the stocks in the sample. The frequency distributions differ notably in the samples with and without zeros, and when comparing Google and Bloomberg search volume series.

Table 2.2: Descriptive statistics of the time series

	Log-ret.	Google Searches		Bloomberg Searches		Trading Volume	
	r_t %	GSV_t	$\log GSV_t$	BSV_t	$\log BSV_t$	TV_t	$\log TV_t$
Whole sample							
Average	0.04	6.80	0.28	2.43	0.23	10,116.96	3.84
Standard deviation	2.67	17.08	0.81	1.27	0.41	9,813.58	0.42
Median	0.05	2.66	0.43	2.00	0.21	7,271.00	3.86
Minimum	−37.61	0.00	−2.56	1.00	−0.43	1.00	0.00
Maximum	53.49	476.28	2.68	4.00	2.20	236,911.00	5.37
Asymmetry	0.46	10.93	−0.63	0.13	0.36	3.31	−0.93
Kurtosis	14.24	188.47	0.04	-1.65	−0.40	23.25	3.34
Averages per stock							
Average	0.05	1.40	0.63	0.92	0.58	9,033.26	3.75
Standard deviation	0.11	1.59	0.49	0.89	0.50	7,092.55	0.39
Median	0.03	1.33	0.59	0.48	0.65	6,855.22	3.75
Minimum	−0.24	0.00	−0.41	0.00	−0.08	671.04	2.58
Maximum	0.37	10.97	1.82	2.67	1.88	40,690.77	4.57
Asymmetry	0.25	4.62	−0.09	0.92	0.24	2.13	−0.80
Kurtosis	2.20	26.11	−0.31	-0.83	−0.50	6.73	1.28

Note: Main statistical measures of the time series of (nominal and adjusted) log-returns, the index that represents the volume of Google search queries made in Brazil, the index that represents the volume of Bloomberg readership and search queries, and the amount of trades, all relative to the stocks that comprised the Ibovespa during the second third of 2018. Search volumes and trading volume are presented in their regular and logarithmic form. The top panel shows measures of the whole sample and the bottom panel shows statistical measures of the averages per stock. We used daily data from 2011 to 2017.

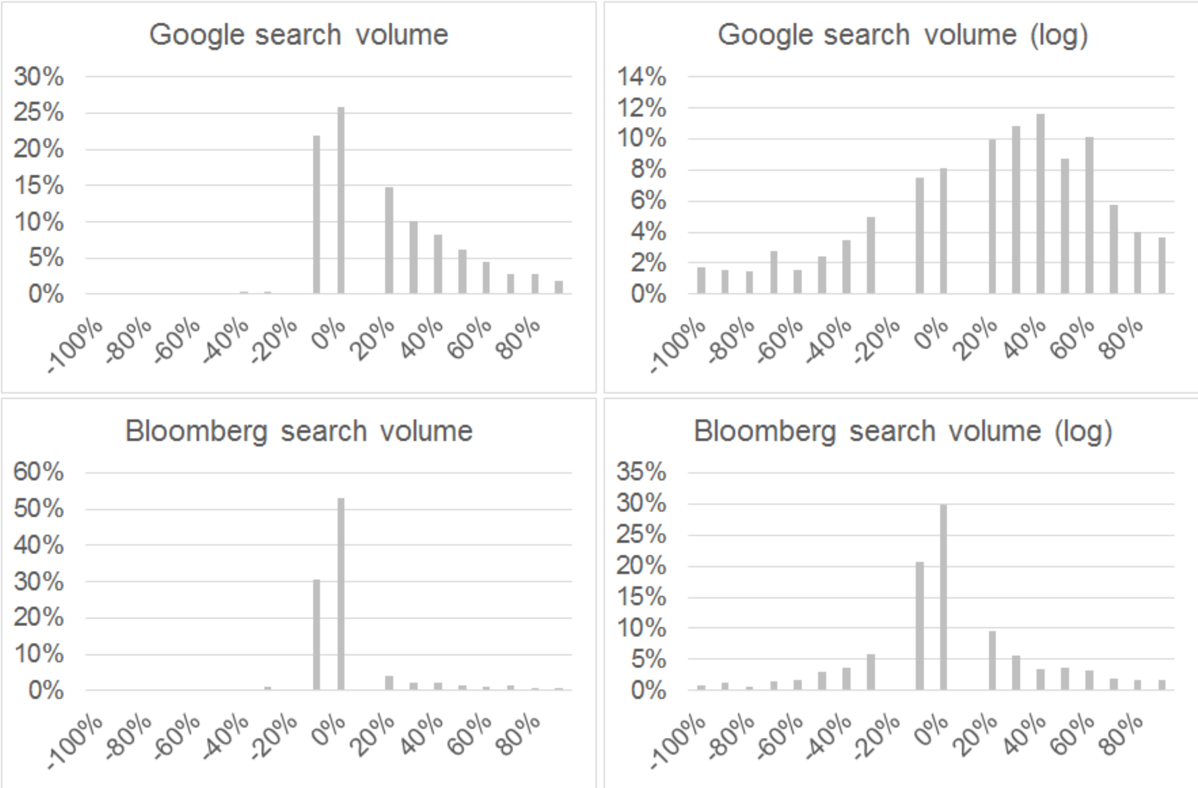

Figure 2.2: Frequency distribution of search volume correlations among stocks
Note: Distribution of the linear correlations of the search volumes of all pairs of stocks. The charts on the top show Google search volume correlations, whereas the charts on the bottom present Bloomberg search volume correlations. The charts on the left show correlations of the series in their regular form, whereas the charts on the right show correlations of the series in their logarithmic form. The stocks are the ones that comprised the Ibovespa during the second third of 2018. Data are in daily frequency, from 2011 to 2017.

On the right side we present the time series in the log scale, discarding observations equal to zero. 73% of the correlations are positive with respect to Google search volume, and well distributed between 20% and 70%. On the other hand, we did not find expressive correlation among Bloomberg search volume time series, however. 51% of the correlations fell within -10% and 10%. Keeping observations equal to zero, negative correlations virtually do not occur anymore.

Correlations between Google and Bloomberg search volume time series for each ticker are shown in Figure 2.3. 98% of the correlations fell within -10% and +10% when we kept observations equal do zero. After discarding these cases, the distribution becomes less concentrated. In this setup, 70% of the correlations fell within -20% and +10%.

At Google and at Bloomberg, search volume index equal to zero may indicate a very short of unavailable amount of searches. Hence, we opted to discard these observations from the sample.

In the next section we present the results of the stock returns modeling through

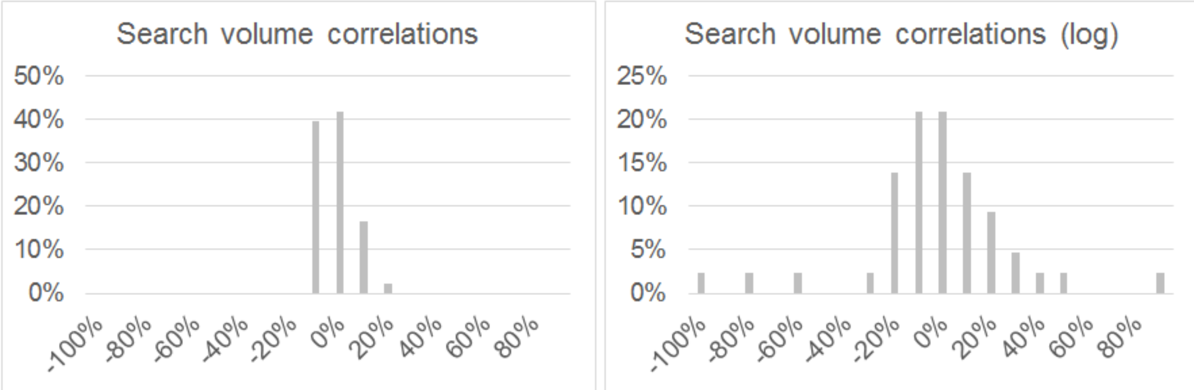

Figure 2.3: Frequency distribution of Bloomberg and Google search volume correlations
Note: Distribution of the linear correlations between Google and Bloomberg search volumes of each stock. The chart on the left shows correlations of the series in their regular form, whereas the chart on the right shows correlations of the series in their logarithmic form. The stocks are the ones that comprised the Ibovespa during the second third of 2018. Data are in daily frequency, from 2011 to 2017.

autoregressive equations. This setup allows the verification of an attention effect on the autocorrelations.

2.3.2 Autoregressive Modeling

Since we have 26 stocks with available Google information and 38 with available Bloomberg information, this section first presents two cases that illustrate the pattern of results and then a table that summarizes the aggregate results for all stocks.

Some simulations were necessary to verify that the results do not expressively change when we add more than two lags. As expected, in most of the cases no return correlation was evidenced. Table 2.3 illustrates an example for the largest Brazilian beverage company, AmBev. None of the regressors (of the returns, their interactions with search volume, or the trading volume) presented significant coefficients in models 1 or 2.

However, we identified autocorrelation in Bloomberg search regressions for 9 stocks and in Google search regressions for 11 stocks, considering two lags. In these cases, we verified whether Internet search volume mitigates these autocorrelations.

Table 2.4 presents the results for a large steel company, Companhia Siderúrgia Nacional. Regressions evidenced significant positive autocorrelations in the first lag considering Google searches and in the second lag regarding Bloomberg searches.

In both cases, the interaction $r \times \log SV$ presented significant negative coefficient in the same lags. In other words, search volume reduced the effect of past returns on current returns. Adding trading volume to the regressions does not affect the results. This finding supports the conjecture that an increase in attention levels reduces return predictability.

Table 2.3: Autoregressive estimation for AmBev (largest Brazilian beverage company)

	Google Searches (GSV_t)		Bloomberg Searches (BSV_t)	
	Model 1	Model 2	Model 1	Model 2
	r_t	r_t	r_t	r_t
r_{t-1}	0.02	0.02	0.05	0.06
	(0.68)	(0.76)	(0.64)	(0.62)
r_{t-2}	0.08	0.08	−0.09	−0.07
	(0.15)	(0.14)	(0.43)	(0.53)
$r_{t-1} \times \log SV_{t-1}$	−0.12	−0.12	−0.12	−0.14
	(0.11)	(0.14)	(0.64)	(0.59)
$r_{t-2} \times \log SV_{t-2}$	−0.05	−0.06	0.29	0.25
	(0.49)	(0.44)	(0.26)	(0.35)
$\log TV_{t-1}$		0.16		0.04
		(0.68)		(0.96)
$\log TV_{t-2}$		0.34		0.47
		(0.39)		(0.55)
Const.	0.02	2.15	0.02	−3.71
	(0.75)	(0.45)	(0.88)	(0.44)

Note: Outcomes of the 2 autoregressive models (Equation 2.1), combining series of log-returns (r_t), their interaction with search volume ($r_t \times SV_t$), and trading volume (TV_t) for AmBev ("ABEV3"):

Model 1: $r_t = c + \sum_{j=1}^{2} [\beta_j r_{t-j} + \delta_j (r_{t-j} \times SV_{t-j})] + \epsilon_t$

Model 2: $r_t = c + \sum_{j=1}^{2} [\beta_j r_{t-j} + \delta_j (r_{t-j} \times SV_{t-j}) + \theta_j TV_{t-j}] + \epsilon_t$

From left to right, the table separately displays outcomes using Google search volume (GSV_t) and using Bloomberg search volume (BSV_t). All series are in daily frequency and in their logarithmic form. We considered the period from 2011 to 2017. The model coefficients of 2 lags are presented with the respective p-values in parenthesis. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Table 2.4: Autoregressive estimation for CSN (large Brazilian steel company)

	Google Searches (GSV_t)		Bloomberg Searches (BSV_t)	
	Model 1	Model 2	Model 1	Model 2
	r_t	r_t	r_t	r_t
r_{t-1}	0.16^{**}	0.16^{**}	0.04	0.05
	(0.01)	(0.02)	(0.64)	(0.55)
r_{t-2}	-0.06	-0.05	0.21^{***}	0.21^{***}
	(0.38)	(0.45)	(0.01)	(0.01)
$r_{t-1} \times \log SV_{t-1}$	-0.16^{**}	-0.15^{*}	0.13	0.11
	(0.04)	(0.06)	(0.41)	(0.49)
$r_{t-2} \times \log SV_{t-2}$	0.00	0.00	-0.34^{**}	-0.35^{**}
	(0.99)	(0.99)	(0.04)	(0.04)
$\log TV_{t-1}$		-0.94		-1.44
		(0.50)		(0.21)
$\log TV_{t-2}$		-1.29		0.81
		(0.40)		(0.51)
Const.	-0.14	-3.00	0.21	-5.97
	(0.49)	(0.67)	(0.23)	(0.12)

Note: Outcomes of the 2 autoregressive models (Equation 2.1), combining series of log-returns (r_t), their interaction with search volume ($r_t \times SV_t$), and trading volume (TV_t) for Companhia Siderúrgica Nacional (“CSNA3”):

Model 1: $r_t = c + \sum_{j=1}^{2} [\beta_j r_{t-j} + \delta_j (r_{t-j} \times SV_{t-j})] + \epsilon_t$

Model 2: $r_t = c + \sum_{j=1}^{2} [\beta_j r_{t-j} + \delta_j (r_{t-j} \times SV_{t-j}) + \theta_j TV_{t-j}] + \epsilon_t$

From left to right, the table separately displays outcomes using Google search volume (GSV_t) and using Bloomberg search volume (BSV_t). All series are in daily frequency and in their logarithmic form. We considered the period from 2011 to 2017. The model coefficients of 2 lags are presented with the respective p-values in parenthesis. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

The summary present in Table 2.5 indicates that the results are in line with our research hypothesis. With respect to Bloomberg searches, out of the 9 significant autocorrelations evidenced in the first or second lag, 8 were accompanied by a reducing effect coming from the interaction of returns and searches. For 6 stocks, this effect was significant at a 10% level.

Regarding Google searches, 9 out of the 11 autocorrelations presented this same mitigating effect, 6 of them statistically significant at a 10% level. This finding indicates that the effect is more salient for Bloomberg searches, suggesting that the attention of professional investors induces less predictability in stock prices than the one of less sophisticated investors.

The effects of the professional and non-professional attention may be correlated. Suppressing the possible influence of this correlation from the results allows a more precise identification of the factors that can explain our findings. In the next section, we perform an additional test to investigate this possibility.

Table 2.5: Summary of the results for the stocks

	Google Searches		Bloomberg Searches	
	1^{st} lag	2^{nd} lag	1^{st} lag	2^{nd} lag
Significant autocorrelation	4	7	4	5
Among those cases, inverted effect of the interaction	4	5	3	5
Among those cases, inverted and significant effect of the interaction	2	4	1	5

Note: Number of cases of each condition, based on the outcomes of the regressions implemented. For each stock, we regressed the returns on lagged returns and interactions between those same lagged returns and Internet search volume. 38 stocks were submitted to the regressions using Bloomberg search volume and 26 were considered in the regressions using Google search volume, after eliminating cases without enough data available. Cases with search volume equal to 0 were disregarded. The observations were converted to their logarithmic form. We adopted two lags in the regressions and a threshold of 10% to consider statistical significance of the coefficients.

2.3.3 Effects of Filtered Non-Professional Attention

The results presented in the previous section suggest that the attention measured by the activity of professional investors affects the market efficiency more substantially than the one indicated by non-professional investors. In this section, we perform an additional analysis to better understand this finding.

The descriptive analysis showed that in many cases there is a correlation between Google and Bloomberg search indices. Therefore, we repeated the regressions replacing the attention measure. Instead of search indices, we used the residuals of the regression of Google searches on Bloomberg searches. Hence, we consider only the Google searches variance that is not explained by the variance of Bloomberg search volume. Equation 2.2

presents these regressions with the residuals ($\eta_{t,i}$). The indices t and i represent time and ticker, respectively.

$$\begin{aligned} GSV_{t,i} &= c_i + \delta_i BSV_{t,i} + \eta_{t,i} \\ rGSV_{t,i} &= \eta_{t,i} \end{aligned} \tag{2.2}$$

This proxy represents the component of non-professional attention that is not correlated with professional attention. If the attention of more sophisticated investors is the one that improves market efficiency, we expect the effects of this proxy to be even lower than the ones found for Google searches in the previous section.

The aggregated results of this specification are presented in Table 2.6. We separately analyzed samples discarding (on the left) and considering (on the right) cases with zero search volume. In the former, there is enough data for only 12 tickers, whereas the latter encompasses 53 tickers. The results correspond to regressions with two lags.

The effect of predictability reduction was evidenced only in few cases. Discarding observations with zero searches, we found autocorrelation in 4 cases. Among them, only in one we verified a statistically significant mitigating effect of the interaction.

Keeping the observations with zero searches, autocorrelation was found in 25 cases. However, only in one of them the attention level significantly reduced this effect.

The results suggest that the professional attention is the one that induces more expressive effects in market efficiency. After removing from non-professional attention the variance explained by professional one, its effect is virtually canceled.

The test results corroborate the hypothesis that behavioral biases give to non-professional investors a role more associated with a volatility inducer than to a market efficiency one.

2.4 Final Considerations of the Chapter

In this chapter, we investigated whether the level of market efficiency is affected by the attention level of investors. The Internet activity level was adopted to measure this attention. Data searching through the Internet is currently a recurring process for keeping informed about publicly-traded assets. Besides, since it is not attached to the stock exchanges, this measure is not an equilibrium outcome, minimizing endogeneity problems.

We used both Google and Bloomberg searches in our study. While the former is popular, free and directed to non-professional investors, the latter is targeted to more sophisticated ones. This allowed us to verify whether attention effects depend on which type of agent is attentive. Previous studies documented behavioral aspects associated with individual investors, such as limited rationality and herd behavior (Chiang & Zheng, 2010;

Table 2.6: Results using $rGSV$ (non-professional attention uncorrelated with professional one)

Using the residuals of the regression of GSV on BSV as attention indicator				
	Without $SV = 0$		With $SV = 0$	
	1^{st} lag	2^{nd} lag	1^{st} lag	2^{nd} lag
Significant autocorrelation	1	3	16	9
Among those cases, inverted effect of the interaction	1	2	7	1
Among those cases, inverted and significant effect of the interaction	0	1	0	1

Note: Number of cases of each condition, based on the outcomes of the additional regression implemented. For each stock, we regressed their Google search volume on their Bloomberg search volume; we then obtained the residuals. After that, we regressed the returns on lagged returns and interactions between those same lagged returns and the residuals of the previous regression. The left part, encompassing 12 stocks, displays results disregarding cases with search volume equal to 0 and considered variables in their logarithmic form. The right part, involving 53 stocks, presents results keeping cases when search volume equals 0 and used the variables in their regular form. We adopted two lags in the regressions and a threshold of 10% to consider statistical significance of the coefficients.

Kahneman, 2003; Sewell, 2007). Since these features have the potential to induce market inefficiencies, our second hypothesis is that professional attention has more potential to induce efficiency than non-professional one.

To perform the analysis, we obtained daily data from 2011 to 2017. This resulted in 1,825 trading days, sufficient for statistical concerns. As far as we search, our work pioneered in such tests at daily frequency. We examined the most material Brazilian individual stocks. Also for the first time, professional and non-professional attention is compared with respect to the effect it induces in the incorporation of information on risky asset prices.

We obtained 67 stocks that were part of Ibovespa, the most popular Brazilian stock index, in the second third of 2018. We opted for examining individual stocks, compared to indices, due to more scarce previous literature regarding the former. Besides, the Bloomberg proxy for information gathering is available only for stocks.

For the analysis, we developed autoregressive specifications of the returns, which included as regressors interactions between those returns and related Internet search indices. In a relevant part of the cases in which autocorrelation was evidenced, these interactions presented an mitigating effect, reducing the autocorrelation level. This result supports our hypothesis that attention induces market efficiency.

Statistically significant effects were found in 67% of the cases based on professional attention and in 55% of the cases based on non-professional one. However, the data presents some level of correlation between those two measures. When we exclude from Google searches the variance explained by Bloomberg searches, only 25% of the results corroborated with our hypothesis. These outcomes contribute to our conjecture that it is the class of professional investors - in general less likely to present behavioral biases (Ben-Rephael et al., 2017) - that induces market efficiency.

The results corroborate with previous studies that describe the information discovery hypothesis (Tantaopas et al., 2016; Vozlyublennaia, 2014). Since we verify that the effect is minimized when non-professional attention is isolated, our findings are also in line with the conjecture of price pressure induction by excess of attention, described in other studies (Barber & Odean, 2008; Da et al., 2011).

We aimed to make some contributions by performing a pioneer analysis of the effect that the attention of different classes of investors may prompt to market predictability. Firstly, we recognize the importance of attention and trading from professional investors to incorporate information to prices. The findings subsidize companies and information providers to define strategies to manage this process of incorporation. The results are even more relevant given the enormous amount of information that people are currently producing and sharing.

Future research may address the implementation of empirical models to better understand the mechanism through which professional attention induces market efficiency. Another promising path to complement our findings is assessing the possibility of gains with investment strategies that account for the higher predictability of stock prices in moments of reduced professional attention.

References

Adachi, Y., Masuda, M., & Takeda, F. (2017). Google search intensity and its relationship to the returns and liquidity of japanese startup stocks. *Pacific-Basin Finance Journal*, *46*, 243–257 (cit. on p. 31).

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31–56 (cit. on pp. 19, 25).

B3. (2015). Metodologia do índice bovespa (cit. on p. 31).

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *The Review of Financial Studies*, *21*(2), 785–818 (cit. on pp. 10, 20, 25, 28, 43, 47, 50, 51, 68, 86, 98).

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307–343 (cit. on p. 19).

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009–3047 (cit. on pp. 10, 15, 43, 88, 97).

Chan, W. S. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, *70*(2), 223–260 (cit. on p. 25).

Chiang, T. C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, *34*(8), 1911–1921 (cit. on p. 41).

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499 (cit. on pp. 10, 11, 20, 43, 47, 48, 52, 53, 56, 57, 60, 65, 66, 68, 75, 87, 97).

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, *52*(3), 1035–1058 (cit. on p. 25).

Falkinger, J. (2008). Limited attention as a scarce resource in information-rich economies. *The Economic Journal*, *118*(532), 1596–1620 (cit. on pp. 19, 26, 27).

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34–105 (cit. on p. 32).

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, *64*(5), 2023–2052 (cit. on pp. 20, 24–26).

Gabaix, X., & Laibson, D. (2005). Bounded rationality and directed cognition. *Harvard University* (cit. on p. 20).

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393–408 (cit. on pp. 19, 21, 22, 51).

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *The Review of Financial Studies*, *17*(2), 439–461 (cit. on pp. 20, 24).

Johansson, E. (2014). Creating daily search volume data from weekly and daily data (cit. on pp. 31, 60).

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, *93*(5), 1449–1475 (cit. on p. 41).

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483–510 (cit. on pp. 19, 22, 23, 51).

Ministério do Planejamento, D. e. G. (2015). Mercado de capital brasileiro está em 12º no ranking mundial (cit. on pp. 11, 30).

Mondria, J., & Wu, T. (2010). The puzzling evolution of the home bias, information processing and financial openness. *Journal of Economic Dynamics and Control*, *34*(5), 875–896 (cit. on p. 20).

Sewell, M. (2007). Behavioural finance. *University of Cambridge*, 1–14 (cit. on p. 42).

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, *96*(2), 158–163 (cit. on pp. 19, 24).

Statista. (2014). Popular online search engines in brazil as of may 2014, based on market share (cit. on p. 31).

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in asia-pacific stock markets. *Pacific-Basin Finance Journal*, *38*, 107–124 (cit. on pp. 29, 30, 32, 43, 51, 57, 58, 98).

Van Nieuwerburgh, S., & Veldkamp, L. (2007). Information immobility and the home bias puzzle. *NBER Working Paper No. 13366* (cit. on p. 20).

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, *41*, 17–35 (cit. on pp. 28, 29, 32, 43, 51).

3 Price Pressure Induction: Effects of the Investor Attention on the Market Volatility

> Never pay the slightest attention to what a company president ever says about his stock.
>
> Bernard Baruch (1870-1965)
> Investor, statesman, political consultant

Abrupt deviations of volatility in the stock market, due to changes in fundamentals, capture the attention of the investor and may lead to a herd behavior. This in turn may generate higher levels of non-fundamental volatility. These dynamics suggest an association between attention measures and the state of the market.

Based on this conjecture, Lux and Marchesi (1999) proposed a volatility model for the stock market that relies on the behavior of fundamentalists and noise traders, acting as inductors of additional volatility.

Previous studies about investor attention did not reject the hypothesis that it induces volatility (Barber & Odean, 2008; Da et al., 2011; Huddart et al., 2009). However, choosing a suitable attention indicator seems crucial in this analysis. Brooks (1998) and Dimpfl and Jank (2016), for instance, did not find significant evidences of predictive power of trading volume on return variance.

Unusual attention levels of individual investors might also lead to implications on the imbalance of orders in the stock market, since, for this class of investors, the range of possibilities to buy is larger than to sell, due to constraints related to short selling (Barber & Odean, 2008). Since long forces imply increase of stock prices, it is reasonable to suspect that there is a relationship between attention measures and abnormal returns.

However, this relationship was not evidenced in the short term by Preis et al. (2010) and Damien, Ahmed, et al. (2013). In general, they did not find significant correlation between search queries and returns, neither superior performance on strategies that considered specific keywords compared to the ones using random keywords.

The central question of this study is: can attention levels help to determine financial market volatility? Particularly, our interest is to verify whether the attention of retail investors can help to anticipate price volatility.

Our results may help market participants to have better conditions to forecast turbulent moments and be better prepared for them. Besides, they can support investors in allocating asset portfolios that are more suitable to their level of risk aversion.

This study contributes to the volatility literature by investigating how attention affects the daily variance of stock returns in Brazil. Also, it allows improvements of predictive models that are essential to asset pricing, risk and portfolio management. We complement the findings of previous studies that used Internet activity to verify how behavioral aspects might influence the stock market (Da et al., 2011; Dimpfl & Jank, 2016).

The rest of this chapter is structured as follows: in the next section, we present previous studies and theoretical frameworks that support our conjecture. After that, we describe the methodological approach and the results for the Brazilian stock market.

3.1 Literature Review

3.1.1 Behavioral Biases in Financial Markets

Biases associated with human behavior have been widely studied by previous research. While some authors used experiments to test the existence of these phenomena, others investigated their implications in capital markets. Burton and Shah (2013) group market anomalies due to behavioral aspects in: the ones derived by the Prospect theory (the reference point, the S-curve and the loss aversion); perception biases (saliency, framing, anchoring and sunk-cost biases); inertial effects (the endowment, the status quo and the disposition effects); some anomalies associated with causality and statistics (the representativeness heuristic, the conjunction fallacy, reading into randomness, the small sample bias and probability neglect); and illusions encompassing talent, skill, superiority and validity.

The effects of psychological factors on market-related decisions have been challenging the traditional view of the economic agent and the efficient market hypothesis. Seminal work on behavioral finance includes papers from Black (1986), about noise trader activity, Benartzi and Thaler (1995), about loss aversion, and Kahneman and Tversky (2013), which proposed the Prospect theory. Hirshleifer (2001) developed a framework that provides a better understanding of the decision biases that affect market prices. Besides, he assessed psychology-based pricing theories that comprise both risk and misvaluation. Moreover, the author described behavioral finance models that considered the impact of heuristics, framing and other insights from neuroscience on different environments.

Hirshleifer (2001) stated that limited attention, processing power and memory are cognitive resource constraints that force individuals to use heuristics for decision-making. He argues that simplifications arising from these heuristics explain most biases identified in experimental psychology.

Based on current literature, the author presented examples of biased decisions resulting from heuristic simplification in financial markets: underestimation of the probabilities of less explicit contingencies due to imperfect recovery of relevant information from memory; self-regulation unconsciously translated into habits due to limited memory, such as consumption out of interests or dividends, not out of principal; mispricing due to favorable evaluation extension of stock characteristics (such as growth or earnings prospects, assumed unrelated); preference to invest in assets with similarities that recall familiar signals, such as local stocks; and mispricing due to utilization of irrelevant but salient cues, which may be abundant in situations of Internet information flood, for instance.

More recent studies typically intend to test and apply these models into a variety of contexts. Frazzini (2006) studied the consequences of the disposition effect in the incorporation of new information to stock prices. This well-documented bias accounts for the tendency of investors to realize gains and ride losses. The author investigated how stock prices responded to positive and negative news. After obtaining data of mutual fund holdings, they constructed a measure based on reference prices and saw that, when the news event and the capital gains (or losses) have the same sign, the post-announcement price drift is stronger. Also, they verified that a strategy based on this "underreaction to news" can be profitable.

Behavioral biases are present in financial markets and leverage uninformed trading. In addition to conventional empirical tests, market modeling may contribute to a broader understanding of how this kind of trading influences general patterns.

3.1.2 Volatility Induction by Noise Traders

The financial market model with different groups of agents, described and simulated by Lux and Marchesi (1999), revealed the effects of the interactions among these groups. The authors classified the participants as fundamentalists, who buy stocks whose prices are lower than their fundamental values (typically the discounted cash flow of the expected future gains), and noise traders, who do not believe in the immediate tendency of price reversion to fundamental values. The group of noise traders, subdivided into optimists and pessimists, seek to graphically analyze the prices and the behavior of the other traders to define tendencies, adopting a typical herd behavior. The model is based on the exogenous source of fundamental values and on the endogenous source of market prices (according to supply and demand balances). Variations in fundamental values were treated as normally distributed stochastic variables, with zero mean and constant variance. Besides, the model considered that, under certain circumstances, agents migrate from one group to another.

The switch between states of pessimism and optimism was treated in the model as a result of two conditions: the current tendency of prices and the prevalent opinion (represented by the difference between the amount of pessimistic and optimistic agents in the market). The shift from a fundamentalist to a noise trader behavior is a consequence of transitory profits of each group: optimistic noise traders profit instantly with positive price chances; on the other hand, the gains for fundamentalists realize in an uncertain future, with a price reversion to the fundamental value; the pessimistic profit with the difference between the returns of alternative investments (assumed constant) and the price changes of the sold asset. Favorable gains for a group change the behavior of participants of the other group.

In the long term, the simulation results in a efficient market with prices reaching fundamental values. However, these prices present different statistical properties, with higher frequency for extreme values and volatility clusters. Absolute returns present long memory and predictability, which is the opposite of what happens for raw returns. Since fundamental values do not present these features, the authors conclude that they come from the interactions among agents, with different strategies and beliefs, and particularly from the migration among groups. Abnormal volatility arises when the number of noise traders is critical, but ends when the prices deviate substantially from their fundamental values (therefore, endogenously). Different parameters lead to similar results and are empirically confirmed.

The proposition of Lux and Marchesi (1999) supports the hypothesis that the participation of noise traders induces additional volatility and substantially contributes to stylized facts of financial markets. Since this participation is preceded by attention, understanding the dynamics of the attention level of this class of investors prompts a better knowledge of market patterns.

3.1.3 Effects of the Attention Level on Financial Markets

Several measures were already used to represent the attention of investors, or, in general, market participants. In some cases, these metrics helped researchers better understand the potentially sub-optimal results of limited attention among different investor classes.

Barber and Odean (2008) considered this approach to analyze transaction data of large and small brokerage firms and money managers of institutional investors. As the purchase of stocks demands more attention from individual investors (considering that short selling is less common for this class of investors), the purpose was to verify whether these investors buy stocks, significantly more than sell, in moments of abnormal attention.

The authors measured attention by three different variables: abnormal trading volume (relative to the past 252 working days), past-day returns and the presence of the ticker in released media news. The analysis encompassed NYSE, American Stock Exchange (Amex) and Nasdaq stock markets.

The results pointed significant imbalances (more buy than sell orders) on orders done by non-professional investors. This pattern was monotonically verified the higher the attention of the investors was. The authors also found evidence that this class of investors did not benefit from this behavior.

Conversely, for institutional money managers, the imbalance was the opposite: more sell than buy orders in periods of higher attention. The behavior of this class of investors tended to follow their predefined strategy.

Following this hypothesis of investor attention proposed by Barber and Odean (2008), Huddart et al. (2009) verified the influence of extreme closing prices on the relative volume and imbalance of orders, which reflect the decision pattern of investors. As extreme prices, they considered the ones that fall out of the range of the last 52 weeks. They collected weekly data of a random sample of 2,000 companies listed at NYSE, Amex and Nasdaq, from 1982 to 2006.

The authors regressed trading volume on dummies that indicated extreme prices, as well as several control variables. As expected, they found that extreme prices determine abnormal amount of transactions. Results were even more pronounced for younger or more volatile companies. After the increase, the amount gradually reduces.

They also found a strong relationship between extreme prices and subsequent returns. The authors based their analysis on the hypothesis that an increase on the number of transactions captures more investor attention. Increased attention would lead to an imbalance towards buying and consequent stock price appreciation. Portfolios formed by smaller or more volatile firms had better price performance.

Hence, the authors divided transactions in two groups based on their size to verify differences between individual and institutional investors. As expected, extreme prices determine abnormal volume and imbalances more expressively among investors with smaller orders.

Other works have analyzed the dynamics between attention and market movements to verify whether more attention induces more efficiency. Tantaopas et al. (2016) named this pressure the information discover hypothesis. They analyzed capital markets of 10 countries from Asia and Oceania. Earlier, Vozlyublennaia (2014) did a similar investigation for security indices of different categories: stocks, bonds and commodities. Such studies rely on the theoretical models of Grossman and Stiglitz (1980) and Merton (1987) about the impact of information on prices. Information is more quickly incorporated into prices under higher attention levels. Besides, to the extent that more data is under the domain of investors, it may result in less information asymmetry, entailing reduced uncertainty and market liquidity.

Variables such as returns or amount of transactions only partially attend the purpose of indicating the intensity of attention, since they result from market equilibria. If attention determines market dynamics, it makes sense to identify proxies that represent attention in a more reasonable way.

3.1.4 Internet Search Queries as an Indicator of Investor Attention

The development of new platforms for the assessment of Internet activity contributed to the identification of potentially more effective measures of the attention level of market participants. Da et al. (2011) investigated the role of Google search queries as a measure of investor attention. They obtained weekly data of the index that represents the amount of searches of the 3,606 individual stocks that comprised the index Russell 3000 from 2004 to 2008. They considered keywords associated with the name of the firm or its major product.

The authors also verified on the Dow Jones archive the frequency of appearance in the news media and articles, classifying by news importance (more or less relevant) and connotation (positive or negative). Moreover, for each firm, they obtained data regarding amount of transactions, release of accounting information, advertising expenses and analysts' coverage.

Interestingly, they did not find significant correlations between Google search queries and all these other proxies for attention level. According to the authors, extreme returns and transaction volumes may be influenced by several other factors. The presence in the news, on the other side, does not guarantee investor attention. The correlation between searches for the ticker and for the main product line of the firm also was not high because the aim of each type of search differs.

The authors implemented a VAR model comprising search queries, transactions volume, absolute abnormal returns and presence in the news. They found that search queries determine all the other variables. Search queries also presented a mean-reversion behavior after abnormal amount of transactions or presence in the news, but persistence after extreme returns.

Regressions of abnormal values of search volume indicated a relevant relationship with firm size, extreme returns and abnormal volume of transactions. The relevance for news depended on their importance. However, in general these variables explained only a minor part of the attention variance.

The authors separately tested covered and marketable limit orders, which are typically done by individual investors. They found significantly positive relationship between search queries and transaction volume, particularly for less informed investors.

They also found evidence that, even controlling for other variables, abnormal search volume determines excess future returns. An interaction term between search queries and market value indicated that the effect is even more expressive for smaller firms, for which buying pressure has greater impact. Increased transaction volume and presence in the news also determined abnormal returns. After three weeks, prices reverted to their previous levels, indicating that fundamental values might not have changed.

During IPOs (Initial Public Offerings), excess of enthusiasm of non-professional investors may lead to abnormal returns on the first day, and reduced performance in the long term. Da et al. (2011) tested the price pressure hypothesis in these cases. Data of 185 offerings from 2004 to 2007 have evidenced significant changes on trading volume around the date of the IPOs. Interestingly, IPOs with higher abnormal search volume presented significantly higher returns on the first day of the offering, better than presence in the news.

Moreover, they found that, in the long term, these stocks had a worse performance. The analysis considered several control variables and portfolio sorting. This long term reversion was not found for firms with high first-day performance but low search volume.

Combined with first-day returns, search queries presented relevant predictive power of long term underperformance. The same was not confirmed for other attention proxies, such as presence in the news, IPO price revision or market sentiment. This suggests that search volume is more likely to capture the attention of individual investors and that this is the one that contributes to high first-day returns, typically reversed in the long run.

One may argue that the expectation of high returns after the IPO may induce abnormal search volume before the IPO. In order to test this anticipation hypothesis, the authors incorporated the market expectancy of first-day returns and verified that, even in this configuration, the predictive power of search queries remained.

Using more recent data, studies present different findings about the relationship between search queries and abnormal returns. Bijl et al. (2016) investigated whether Google search queries predict stock returns on a weekly basis. They obtained prices of 431 stocks that comprised the S&P 500, from 2008 to 2013. After calculating excess returns, normalized search volume, excess trading volume and realized volatility, they evidenced that higher search volume lead to negative returns in the short term. Panel regressions were used combining 5 lags for the independent variables as well as dummies indicating the month of January and the global financial crisis, and lagged interaction variables.

The results were even more expressive during the global financial crisis. They found that search volume determined returns more than trading volume did. The findings were robust to different configurations. Using random keywords for search queries or lagged search volume, the results did not persist.

Hence, they implemented a trading strategy based on buying stocks with low search volume. They formed weekly balanced portfolios over a 5-year period. The performance was 16% higher than the one of a simple equally-weighted portfolio. However, considering transaction costs, this active strategy presented a performance 5% lower than the benchmark.

The predictive power of search volume on volatility was also the object of investigation in recent studies. Dimpfl and Jank (2016) analyzed the effect of Google search queries on price volatility in the American stock market. They obtained data from 2006 to 2011 of the Dow Jones index, which tends to capture more attention of individual investors, compared to other stock exchanges. Realized daily volatility was obtained through intraday returns, using a 10-minute interval.

The authors examined the index that represented the volume of search queries of the keyword “dow” done in the United States in negotiation days. They concluded that this keyword is more used by investors interested in the stock index.

Time series of realized volatility and search volume presented a correlation of 82% in the period. Trading volume correlated with search queries by 57% and with volatility by 57%.

The VAR model accounted for daily logarithmic series of realized volatility, search volume and trading volume, with a maximum lag of 3 days. As expected, all the variables had significant autocorrelation for almost all lags. Besides, results showed that only current — not past — volatility determines search volume.

The bivariate models indicated that last day higher search volume significantly determines — and Granger-causes — higher volatility. On the other side, the turnover only marginally determines search volume, while, in the opposite direction, the results were more expressive. Granger causality tests confirmed this findings.

Results persisted for the trivariate model: search queries significantly determined volatility, while trading volume did not. Besides, volatility did not predict search volume. The Granger causality confirmed that search volume — and not turnover — determined volatility and that trading and search volume had a reciprocal predictive power.

Patterns were verified also by impulse response functions between search queries and volatility. The authors treated contemporaneous fundamental volatility as exogenous, with unilateral and immediate effect over search volume. As expected, both variables strongly responded to their own shocks. The persistence was evidenced by a low decaying function. Besides, results were particularly noteworthy in the response of volatility for a shock in search queries, which lasted for around 60 trading days.

In order to test the ability to predict, Dimpfl and Jank (2016) made in- and out-of sample tests, for different horizons and frequencies of both volatility and search volume. In addition to traditional autoregressive models, the authors tested the Heterogeneous Autoregressive (HAR) model, which better captures long memory aspects. Bivariate models (VAR and VHAR) were compared to univariate ones (AR and HAR). The predicting performance of volatility was measure by the mean squared error and the quasi-likelihood loss functions, as well as comparing the R^2 of the model with predicted and realized values.

Regarding the univariate sets, in-sample results of the AR(4) model were more accurate. Interestingly, the bivariate sets presented better results for all performance measures.

Out-of-sample predictions encompassed one-day, two-week and three-week scenarios. Initially, the model was estimated using a sample of two years and re-estimating in each moment using all past available information. Considering univariate sets, results pointed that the HAR model had superior prediction power in higher horizons. Bivariate sets performed generally better for considering the effect of the attention of individual investors. For both in-sample and out-of-sample predictions, the VHAR model performed better than the HAR model especially in moments of higher turbulence. Results persisted even considering the 1-day delay that Google Trends takes to disclose updated search volume index.

The authors verified whether volatility anticipation using search queries may lead to economic gains. The simulation was based on a quadratic utility function (with risk aversion coefficients of 1, 5 and 10) and a variance target of 12%. Portfolio gains were realized and prices were recalculated at each interval. Short selling was not allowed and all borrowings were at the risk-free rate, using United States (US) sovereign funds as benchmarks.

Two different scenarios were designed: one with weights calculated based on the predicted variance of the HAR model, and the other one based on the fitted values of the VHAR model. These strategies were compared to a buy & hold portfolio with 50% risky assets. The authors compared different strategies by measuring the maximum performance fee that a investor would accept to migrate from one portfolio to another one, keeping the original utility. For a risk aversion coefficient of 1, the investor would accept a fee of 112 basis points to migrate from the buy & hold portfolio to one based on the HAR model. However, 121 basis points would be accepted to switch to a strategy based on the bivariate model, resulting in a gain of 8.99 basis points per year. This pattern persisted for higher risk aversion coefficients.

While previous studies confirmed that search queries are indeed proxies for investor attention, we believe that there is still a lot to be investigated in financial markets using this kind of modern indicator. Individual investors represent a relevant kind of market agent, with particular attributes that influence their trading behavior. Hence, in this study, we focused on verifying whether the attention of these agents induces market volatility.

3.1.5 Empirical Evidence about Investor Attention in Brazil

Few recent studies have addressed the relationship between investor attention and financial market movements in Brazil. De Souza et al. (2017), for instance, verified whether media coverage determines increased trading volume considering the market performance on the previous day.

Media coverage was measured in hand-collected daily news in printed versions of highly circulated national newspapers. After collected, news were filtered and some were classified as good, bad or not relevant. Their sample comprised daily data from 2010 to 2015 of 17 stocks of Brazilian firms accounted in the Ibovespa. Firms were divided in 2 groups, according to size and liquidity. A dummy variable to indicate good periods (when the Ibovespa was among the 10% highest values of the last 2 years) and past returns were also included in the model. When there is a shock in news, they found that volume increases on good days and decreases on the next day. Moreover, the market took longer to react to negative news and the effects were more visible for the group of smaller or less liquid firms.

Internet activity was used as a proxy for investor attention in the Brazilian market only very recently. Rodolfo et al. (2017) investigated whether search queries determine financial trading volume, turnover and abnormal returns in the Brazilian stock market. The sample comprised series of 232 public firms from 2010 to 2015 in a weekly basis. Abnormal returns were calculated using the Capital Asset Pricing Model (CAPM) and the Ibovespa as references. They also considered the impact on absolute abnormal returns.

They found low correlations among proxies of attention levels, as Da et al. (2011) did. The VAR model indicated that current investor attention determines more abnormal returns and turnover in the next weeks. However, when analyzing the opposite effect, the authors found only a significant effect, although negative, of lagged absolute abnormal returns on attention.

After that, they implemented regressions to verify whether price movements are transient due to limited attention. The authors used panel regressions with fixed effects and dynamic panel regressions estimated by Generalized Method of Moments (GMM) with differentiated variables. Results evidenced that investor attention influences market activity, but does not drive temporary price pressures. In other words, Rodolfo et al. (2017) did not find that the contemporaneous impact of attention on prices was later reversed. The authors attributed the findings to the weekly frequency, which might not truly capture this kind of price dynamics. Besides, they associated the results with local specificities, such as less information to process, less representativeness of retail investors, or even large influence of external factors.

Also using weekly Google search queries, Miragaya (2017) developed an event study to check whether abnormal returns lead to higher investor attention. The sample encompassed 58 firms that composed the Ibovespa from 2007 to 2016. Different from Rodolfo et al. (2017), the author found that absolute abnormal returns determine more attention in the next week. Besides, an asymmetry pattern was found, meaning that negative returns induce more attention than positive ones.

Miragaya (2017) argued that it happened possibly because buying opportunities capture more attention than selling ones. Also, the disposition affect determines that people are reluctant to sell losing stocks in their portfolio. So, these stocks would require more analysis (and attention) than winning ones for the trading decision. The effect of returns on attention lasted for 3 weeks. Besides, trading volume also aroused attention. The methodology considered linear regressions with control variables and corrections for serial correlations.

Despite improving our knowledge with respect to the effect of investor attention in the Brazilian context, these approaches did not always have converging results. Moreover, we believe that there is still much to know about the implications of variations in the degrees of attention coming from individual investors. We addressed this issue in our study.

3.1.6 Additional Considerations Regarding Attention and Financial Markets

Once available, any new information can be received in different ways (depending on the primarily level of attention) and in different moments (or even not be received) by each investor. This is one of the major advantages of using Internet search volume as measure of investor attention. It indicates the moment when the information is sought, not only available, and correlates with other attention measures (Da et al., 2011).

Another important aspect is that, after having access to the information, agents may or may not trade (or participate, more generally) in stock exchanges. In other words, participation is preceded by attention, but the opposite is not necessarily true. Da et al. (2011), however, found a strong and direct connection between search volume and trading volume of individual investors. This relationship supports the hypotheses we intend to explore. They base on the assumption that, in moments of higher attention, a significant portion of noise traders will induce volatility to stock markets. A proper comprehension of the effects of the attention levels of this class of investors allows more appropriate and optimized market strategies.

The development level of the stock market and of the economy, and even cultural aspects, may influence the relationship between attention and volatility. The cross-country analysis of Tantaopas et al. (2016) considered these explanations. Four of the ten countries from Asia and Oceania analyzed are considered emergent ones and the rest of them are developed ones. The authors investigated the relationship between investor attention and market variables (returns, volatilities and trading volume), using indices that represent the capital market of each country.

Tantaopas et al. (2016) stated differences among countries may be explained by factors such as: less representativeness of individual investors, differences between Western and Eastern cultures, and more delay in the market response (or worse information transmission) of less developed economies. Compared to more developed ones, the Brazilian stock market is smaller, less diversified with respect to firms and segments and has a smaller proportion of individual investors and more limited legal protection (de Ávila Marques et al., 2015). These specificities imply differences in results when we compare to the ones found in other countries.

Empirical evidence of particularities in developed financial markets and effects of the attention of agents justify an examination in the context of emerging countries. We opt in this study to look for evidence of the influence of individual attention on the dynamics of the Brazilian financial market. We state that individual (potentially noise) traders will induce more volatility to the stock market.

The diagram presented in Figure 3.1 is similar to the one described in Figure 2.1. However, now we restrict the analysis to non-professional investors, who, acting as attentive noise traders, induce non-fundamental volatility to the markets, endogenously, by their interactions, by transitions between states of optimism and pessimism and by heterogeneity of beliefs. These interactions give prices remarkable statistical properties, such as volatility clusters and a higher frequency of extreme values.

Following theoretical and empirical research, we established the following hypothesis in the Brazilian stock market:

H_1: higher attention of individual investors induces volatility in the stock market

In the next sections, we present our methodological approach and findings.

3.2 Methodology

In this part of the chapter, we describe how the hypothesis of attention-induced volatility was tested in the Brazilian stock market. We detail the sample data, the source of information, empirical models, interest variables and expected relationships. Empirical studies examined this effect in developed countries, such as the one from Dimpfl and Jank (2016).

3.2.1 Data Description

For this test we obtain data from the Ibovespa. It is the most popular index of the Brazilian stock exchange, comprising the tickers that are more representative of the local market. The search for market indices tends to be less ambiguous than the one from

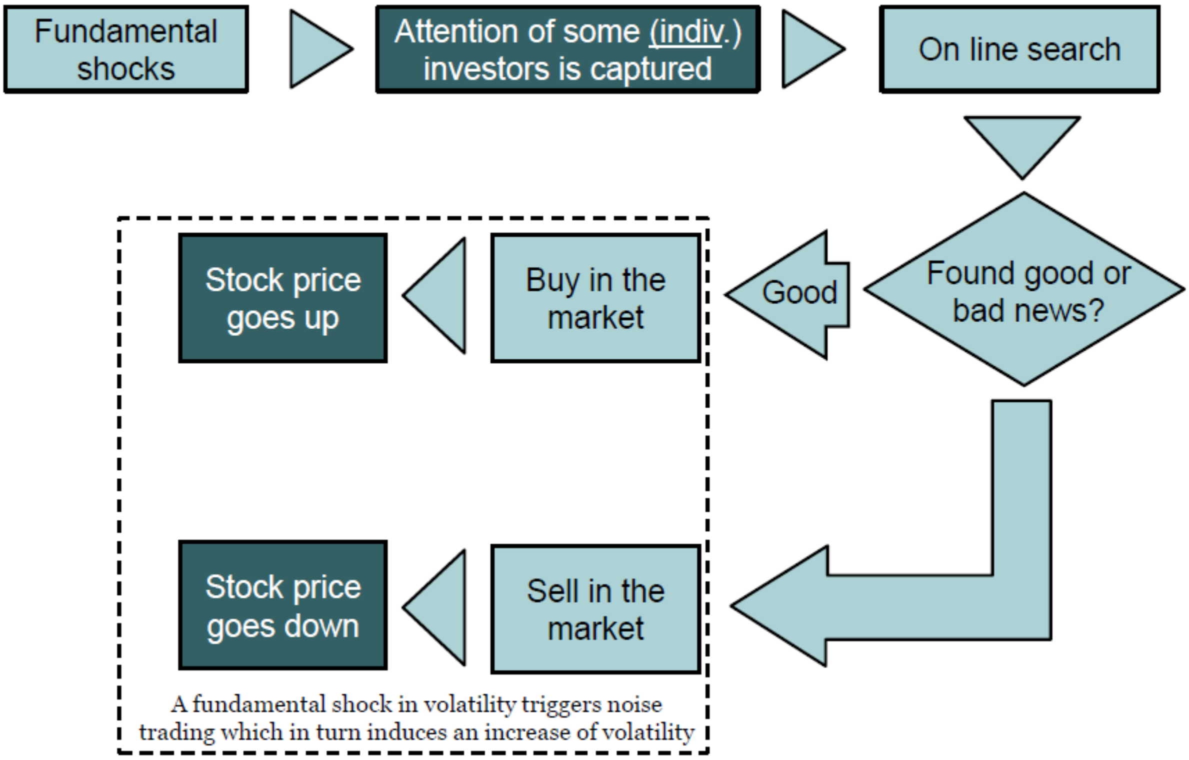

Figure 3.1: Flow diagram of the role of non-professional attention in the volatility of prices

Note: Based on the theories of Dimpfl and Jank (2016) and Lux and Marchesi (1999).

individual stocks, which favors our choice for the forementioned experiment (Da et al., 2011). Besides, forward and future markets of the Brazilian stock exchange (especially Ibovespa futures called "minicontracts") have larger share of individual investors when compared to the spot market (BMFBovespa, 2017). More information about this measure can be obtained in Chapter 1.

We obtained daily nominal log-returns and trading volume relative to the Ibovespa, in the software Economática, from 2004 to the first quarter of 2016. The returns are adjusted for dividends.

At Google Trends, we had access to daily time series of the index that represents the relative search volume of the keyword "Ibovespa" at Google engine. We considered only searches performed in Brazil, in order to minimize effects associated with timezones. We understand that the period is long enough to contain either calm and turbulent periods. Moreover, Google Trends series before 2004 are not available.

We considered only trading days. Each quarter was obtained separately. Then, we concatenated quarters using the method proposed by Johansson (2014). We normalized the series to obtain an unitary mean.

3.2.2 Modeling

Volatility models of stock returns are studied since they were adopted as relevant risk measure. AutoRegressive Conditional Heteroskedasticity (ARCH) models were initially proposed by R. F. Engle (1982) and extended by Bollerslev (1986). He proposed the Generalized ARCH (GARCH), a more comprehensive version that gained relative popularity. These models considered that volatility varies with time and can be predicted using realized values.

Hence, we adopted this method to model the conditional volatility of Ibovespa (σ_t^2). It is represented by the following equations:

$$r_t = \mu + a_t \tag{3.1}$$

$$a_t = \sigma_t \epsilon_t \tag{3.2}$$

$$\sigma_t^2 = \alpha_0 + \sum_{t=1}^{s} \alpha_i a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2 \tag{3.3}$$

The model assumes no serial correlation on the returns r_t, which is tested using the Ljung-Box method and information criteria. The series $a_t = \mu - r_t$ corresponds to the shock in time t, with finite unconditional variance. The component ϵ_t is a random independent variable and identically distributed, with zero mean and unitary variance.

The orders s and m indicated the lags of the components of the volatility model. We find more adequate orders and distribution of parameters by analyzing the significance of the coefficients, the results of the test Lagrange Multiplier test (LM) and Ljung-Box tests, residuals, Bayesian Information Criteria (BIC) and Akaike Information Criteria (AIC). We expected that the coefficients α_i and β_j be non-negative and statistically significant.

After that, we analyze the relationship between volatility and search volume by modeling autoregressive vectors (VAR). We use information criteria to define the most proper maximum lag p. Considering $\mathbf{x}_t$ a vector containing volatility time series (σ_t^2), Google search volume (GSV_t) and trading volume (TV_t), all of them in the logarithmic form, the $VAR(p)$ model is presented as follows:

$$\mathbf{x}_t = \mathbf{c} + \sum_{j=1}^{p} \beta_j \mathbf{x}_{t-j} + \boldsymbol{\epsilon}_t \tag{3.4}$$

The component $\mathbf{c}$ is the vector of constants and $\boldsymbol{\epsilon}_t$ is the vector of shocks, i.i.d. We built three different models, combining three series: model 1 considers volatility and search volume series; model 2 encompasses volatility and trading volume time series; model 3 accounts for the three series. We expect the search volume coefficients to be positive and significant in the volatility model, even considering the trading volume as an additional regressor. Trading volume will be replaced by financial trading volume (FTV_t) to check the persistence of the outcomes. We will confirm the expected relationships using the Granger causality test.

3.3 Empirical Results

In this part of the chapter, we describe details and outcomes of our investigation. A set of methods were applied to verify whether more individual investor attention induces additional volatility of stock prices in the context of the Brazilian stock exchange. Individual attention was measured by Google search volume of associated keywords.

3.3.1 Descriptive Analysis

Search volume data were collected between 3 p.m. and 6 p.m. of 06/18/2016. The keyword "Ibovespa" was chosen because it has more queries than similar ones, such as "Ibov", and because it refers specifically to the index. "Bovespa", for instance, may refer to the name of the stock exchange until it was changed to B3. The correlation between the keyword "Ibovespa" and similar ones is shown on Table 3.1.

Table 3.1: Correlation between "Ibovespa"-related keywords

Keyword	Correlation	Representativeness
Bovespa	0.9570	6.25
Ibov	0.7939	0.04
Ações	0.7831	8.17
Cotação	0.7829	14.67
BMF	0.5289	0.09

Note: Correlations and proportion of unique and correlated keywords with respect to "Ibovespa", performed in Brazil since 2004, obtained from Google Correlate and Google Trends. Data extracted at 10 p.m. on 07/23/2016.

We disregarded data from the 11th to the 24th of April, 2014, since they indicated no search queries, which we assumed was a flaw. Table 3.2 displays descriptive statistics of the series used in this analysis, in its regular and logarithmic form. Each series encompasses 3,018 observations.

The table evidences the high variability of the series of returns and search queries, as well as the substantial decrease in asymmetry and kurtosis of search queries after converting them to the logarithmic form.

Table 3.2: Descriptive statistics of the time series

	Log-returns	Google search volume		Trading volume		Financial trading volume	
	r_t %	GSV_t	$\log GSV_t$	TV_t Million	$\log TV_t$	FTV_t BRL Billion	$\log FTV_t$
Average	0.03	1.23	-0.10	0.38	0.67	4.60	0.58
Standard deviation	1.79	2.15	0.37	0.33	0.54	2.44	0.30
Median	0.06	0.74	-0.13	0.30	-0.53	4.88	0.69
Minimum	-12.10	0.07	-1.16	0.01	-2.10	0.20	-0.70
Maximum	13.68	54.43	1.74	2.41	0.38	25.90	1.41
Asymmetry	-0.01	12.02	0.51	0.76	-0.48	0.75	-0.96
Kurtosis	4.95	219.54	0.92	0.11	-1.13	-0.96	0.17

Note: Main statistical measures of the time series of (nominal and adjusted) log-returns, volume of search queries made in Brazil (Google Trends index), trading volume (in amount and value of trades), all relative to the Ibovespa. Search and trading volume are presented in their regular and logarithmic form. We used daily data from 2004 to the first quarter of 2016.

3.3.2 GARCH Modeling

Firstly, we used daily series of log-returns of Ibovespa to assess its conditional volatility. The t-test of the mean did not reject the null hypothesis (p-value 38.63%). Although the Ljung-Box test rejected the hypothesis of no autocorrelation (p-value 0.19%), the

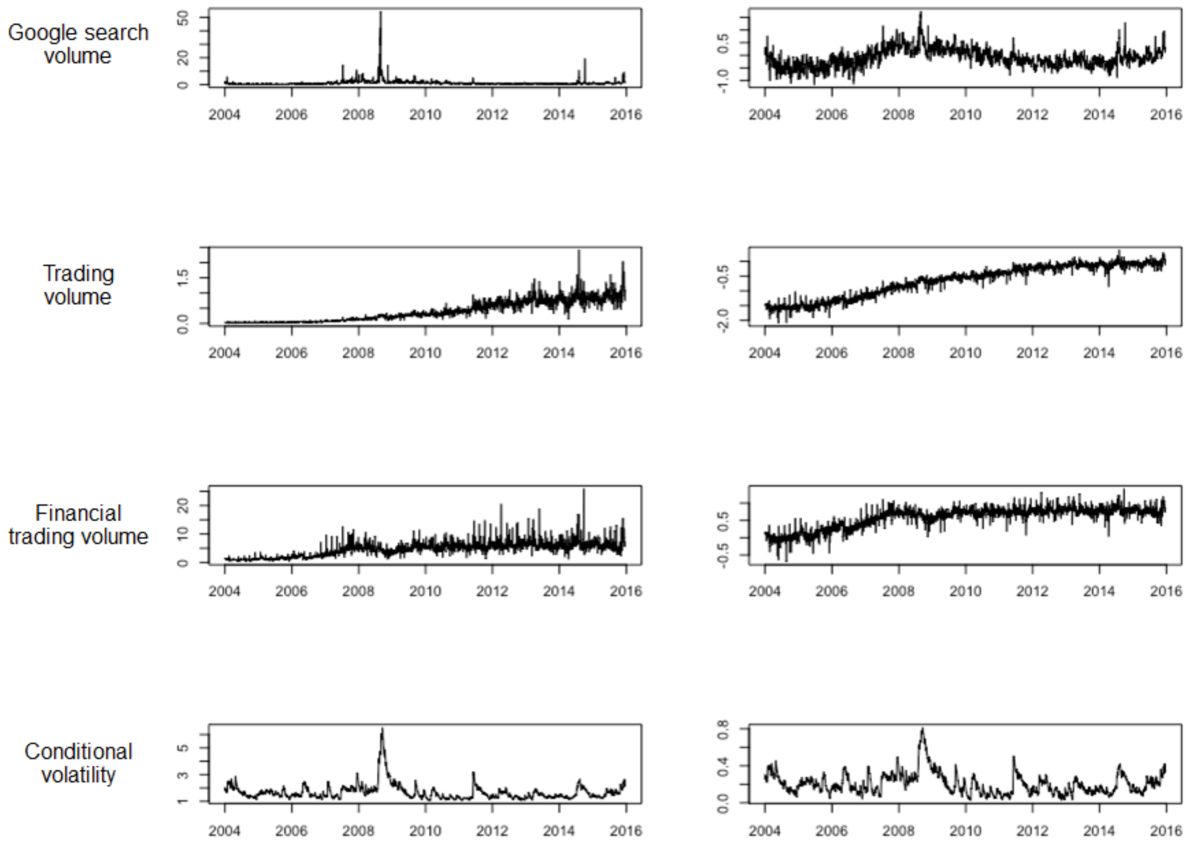

Figure 3.2: Graphical visualization of the series

Note: Evolution of search volume (Google Trends index), the trading volume (in million), the financial trading volume (in BRL billion) and the conditional volatility, measured according to a GARCH(2,1) t-distributed model, with respect to the Ibovespa, from 2004 to the first quarter of 2016. The regular form is shown on the left side and the logarithmic form on the right side.

information criteria statistic indicated null orders as more adequate for modeling the mean (Equation 3.1).

As expected, the Ljung-Box test of the squared series and the Lagrange multiplier test (LM) indicated the presence of ARCH effects (p-value below 0.00%). The GARCH models (Equation 3.3) of orders (2,1) and (2,2), with either t or normal distributions, were considered adequate according to LM and Ljung-Box tests of the residuals, as we show in Table 3.3. Higher-order models presented non-significant coefficients. We selected the t-distributed model with order (2,1), considering information criteria statistics and parsimony.

In Figure 3.2, we show graphically, in normal and logarithmic forms, the four series involved in our autoregressive vectors modeling: the volatility based on a GARCH(2,1) model with t distribution, the Internet search queries, the trading volume and the financial trading volume, all associated with the index. The plots of volatility and search volume evidence the peak related to the 2008 world financial crisis.

Table 3.3: GARCH modeling of the Ibovespa log-returns

	GARCH(2,1)		GARCH(2,2)	
	Normal	t	Normal	t
α_0	0.05*	0.06**	0.05**	0.06**
ω	0.07***	0.07***	0.10***	0.08***
α_1	0.03*	0.02	0.03*	0.02
α_2	0.05***	0.06***	0.09***	0.08***
β_1	0.90***	0.90***	0.36	0.57*
β_2			0.49*	0.30
Distrib.		10.00***		10.00***
Log likelihood	-5,695.86	-5,678.64	-5,694.18	-5,678.13
Ljung-Box tests				
R				
Q(10)	9.90	9.85	10.03	9.89
Q(15)	14.20	14.11	14.17	14.06
Q(20)	24.50	24.44	24.06	24.15
R^2				
Q(10)	12.07	12.79	9.97	10.44
Q(15)	18.85	19.82	16.51	17.24
Q(20)	25.95	26.86	23.53	24.18
LM Test	14.01	14.78	11.51	12.13
Information Criteria				
AIC	3,778	3,767	3,777	3,767
BIC	3,788	3,779	3,789	3,781
SIC	3,778	3,767	3,777	3,767
HQIC	3,781	3,771	3,782	3,772

Note: GARCH modeling summary of the daily log-returns of the Ibovespa (nominal and dividend adjusted), from 2004 to the first quarter of 2016. The model is presented in Equation 3.3:

$$\sigma_t^2 = \alpha_0 + \sum_{t=1}^{s} \alpha_i a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}^2$$

We selected the orders (2,1) and (2,2), with normal and t distributions, since they were considered more adequate according to the Ljung-Box and LM tests. *, ** and *** denote significance at 10%, 5% and 1%, respectively.

Table 3.4 shows the correlations among the series, in their regular and logarithmic forms. As expected, the correlation between the amount of orders (trading volume) and their monetary value (financial trading volume) is relatively high and positive (varying from 0.73 to 0.90).

It is also clear that search volume does not correlate significantly (varying from -0.03 to 0.39) with trading volume or financial trading volume. As Da et al. (2011) mentioned, one of the reasons might be that trading volume results from an equilibrium of several economic factors, besides the attention level of the agents.

The Ibovespa volatility presents reasonable correlation (varying from 0.41 to 0.46) with search volume. This suggests that attention immediately arouses the occurrence of a volatility shock, a particularity also evidenced by Dimpfl and Jank (2016). On the other side, the volatility shows relatively low (and negative, varying from -0.15 to -0.07) correlation with trading volume and financial trading volume.

Table 3.4: Correlation matrix of the series

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
a. Google search volume (GSV_t)	1							
b. $\log GSV_t$	0.67	1						
c. Trading volume (TV_t)	-0.03	0.04	1					
d. $\log TV_t$	0.06	0.22	0.90	1				
e. Financial trading volume (FTV_t)	0.14	0.31	0.77	0.83	1			
f. $\log FTV_t$	0.16	0.39	0.73	0.90	0.92	1		
g. Conditional volatility (σ_t^2)	0.46	0.43	-0.13	-0.09	-0.08	-0.07	1	
h. $\log \sigma_t^2$	0.41	0.42	-0.15	-0.13	-0.11	-0.10	0.96	1

Note: Sample contemporaneous correlations between the daily series, from 2004 to the first quarter of 2016, of Google search volume, trading (financial) volume and Ibovespa conditional volatility, measured by a GARCH(2,1) t-distributed model.

3.3.3 VAR Modeling

Table 3.5 presents the result of the autoregressive modeling, as shown in Equation 3.4, resulting from the logarithmic form of the volatility, search volume and trading volume series. Specifications 1 and 2 combine two of the three series, while specification 3 encompasses the three series.

Using BIC and AIC as information criteria, models $VAR(6)$ and $VAR(10)$ presented more adequate orders. We opted for the model with a maximum lag of 6 days. It is more parsimonious, considering that the results of the $VAR(10)$ modeling were not qualitatively different. Panel A of Table 3.5 shows the estimates and p-values of each coefficient.

As expected, the results show that all series present strong serial correlation. In all specifications, the dependent variable presented significant effects in most of its lagged components. The effects are always positive in the series of trading volume. For volatility and search queries, the coefficients showed different signs depending on the lag.

In the volatility equation ($\log \sigma_t^2$), even after considering the effect of serial correlation, search volume for the previous day ($\log GSV_{t-1}$) entered significantly at 1%, in both the configurations 3.1 and 3.3. This corroborates with the hypothesis that the attention level predicts higher price fluctuations. In the third lag, the coefficient presents significant negative value, indicating that the effect is mitigated after some period.

Moreover, in the search volume equation ($\log GSV_t$) we do not verify a contribution of the volatility components. We did not find statistically significant coefficients for $\log \sigma_t^2$ in any lag, in models 1 and 3. This result indicates that the additional volatility induced by attention is more pronounced than an effect on attention coming from volatility. This finding was also verified by Dimpfl and Jank (2016) in the US stock market.

The inclusion of trading volume ($\log TV_t$) in the model 3 does not alter the relationship between searches and volatility. Besides, $\log TV_t$ of the last day presented positive coefficient (statistically significant at 1%) in the search volume equation (models 2 and 3). In the third lag, the coefficient is negative and statistically significant at 1%.

The coefficient that indicates that search queries are relevant in the model of trading volume was also statistically significant at 1% in the first lag. There results demonstrate that increased attention of noise traders to the stock market anticipates more trading volume, which, in turn, calls more attention (for a while). This evidence was also verified by Da et al. (2011) in the American market.

The model 3 indicated that $\log TV_t$ explains part of the variation in $\log \sigma_t^2$), with positive coefficients in more recent lags and negative ones in the other days. We also found that volatility determines trading volume, but in a less expressive level, with less significant components.

The Granger causality test statistics are presented in the panel B of Table 3.5. The results evidence that current search volume Granger-causes volatility tomorrow, which is coherent with the coefficients found in the panel A and with the research hypothesis. The tests also indicate that the trading volume in previous days provides relevant information that captures current attention (and search volume) and volatility. On the other side, shocks in the attention level directly determine future trading volume. These relationships were found both in models 1 and 2 (bivariate) and in model 3 (trivariate). As well as in the coefficient analysis, we did not find conclusive evidence that past volatility Granger-causes noise traders attention, despite determining trading volume.

We redid the modeling using the financial trading volume instead of the trading volume. All relationships found between Ibovespa search volume, transaction volume and volatility persisted, except for few cases in higher lags.

In Figure 3.3, we provide the impulse response functions in the bivariate model involving volatility and search queries (3.1). The charts (a) and (d) confirm the persistent

Table 3.5: Vector autoregressive modeling

Panel A: Vector autoregressive estimation

	Model 1		Model 2		Model 3		
	$\log \sigma_t^2$	$\log GSV_t$	$\log GSV_t$	$\log TV_t$	$\log \sigma_t^2$	$\log GSV_t$	$\log TV_t$
$\log \sigma_{t-1}^2$	1.38***	-0.28			1.36***	-0.11	-0.04
	(0.00)	(0.18)			(0.00)	(0.61)	(0.73)
$\log \sigma_{t-2}^2$	-0.55***	0.29			-0.51***	0.18	-0.14
	(0.00)	(0.41)			(0.00)	(0.63)	(0.44)
$\log \sigma_{t-3}^2$	0.23***	-0.03			0.22***	-0.05	-0.33*
	(0.00)	(0.95)			(0.00)	(0.90)	(0.10)
$\log \sigma_{t-4}^2$	-0.13***	0.43			-0.10***	0.44	0.32*
	(0.00)	(0.25)			(0.00)	(0.24)	(0.10)
$\log \sigma_{t-5}^2$	0.06**	-0.54			0.04	-0.59*	-0.01
	(0.04)	(0.12)			(0.15)	(0.09)	(0.97)
$\log \sigma_{t-6}^2$	-0.02	0.19			-0.03	0.22	0.14
	(0.36)	(0.36)			(0.11)	(0.28)	(0.18)
$\log GSV_{t-1}$	0.01***	0.64***	0.64***	0.03***	0.01***	0.63***	0.03***
	0	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\log GSV_{t-2}$	0.00	0.18***	0.18***	0.00	0.00	0.18***	0.00
	(0.16)	(0.00)	(0.00)	(0.79)	(0.39)	(0.00)	(0.72)
$\log GSV_{t-3}$	-0.01***	0.04*	0.05**	0.00	-0.01***	0.05**	0.00
	(0.00)	(0.06)	(0.03)	(0.84)	(0.01)	(0.04)	(0.99)
$\log GSV_{t-4}$	0.00	-0.03	-0.02	-0.02**	0.00	-0.03	-0.02*
	(0.68)	(0.25)	(0.29)	(0.04)	(0.49)	(0.23)	(0.10)
$\log GSV_{t-5}$	0.00	-0.09***	-0.09***	0.00	0.00*	-0.09***	0.00
	(0.23)	(0.00)	(0.00)	(0.83)	(0.08)	(0.00)	(0.96)
$\log GSV_{t-6}$	0.00	0.16***	0.16***	-0.01	0.00	0.16***	-0.01
	(0.28)	(0.00)	(0.00)	(0.54)	(0.49)	(0.00)	(0.61)
$\log TV_{t-1}$			0.10***	0.46***	0.03***	0.12***	0.44***
			(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\log TV_{t-2}$			-0.02	0.12***	0.03***	-0.01	0.11***
			(0.59)	(0.00)	(0.00)	(0.75)	(0.00)
$\log TV_{t-3}$			-0.15***	0.04**	-0.04***	-0.14***	0.04**
			(0.00)	(0.04)	(0.00)	(0.00)	(0.05)
$\log TV_{t-4}$			0.06	0.14***	0.00	0.05	0.16***
			(0.14)	(0.00)	(0.40)	(0.18)	(0.00)
$\log TV_{t-5}$			-0.01	0.16***	-0.02***	-0.02	0.17***
			(0.88)	(0.00)	(0.00)	(0.68)	(0.00)
$\log TV_{t-6}$			0.02	0.08***	-0.01***	0.02	0.07***
			(0.50)	(0.00)	(0.01)	(0.69)	(0.00)
Const.	0.01***	-0.02***	0.00	0.00	0.00***	-0.02**	0.01**
	(0.00)	(0.01)	(0.79)	(0.32)	(0.00)	(0.02)	(0.04)

Panel B: Granger causality tests

	Model 1		Model 2		Model 3		
	$\log \sigma_t^2$	$\log GSV_t$	$\log GSV_t$	$\log TV_t$	$\log \sigma_t^2$	$\log GSV_t$	$\log TV_t$
$\log \sigma_{t-j}^2$		1.65				1.93*	8.73***
		(0.13)				(0.07)	(0.00)
$\log GSV_{t-j}$	10.95***			3.43***	7.22***		3.96***
	(0.00)			(0.00)	(0.00)		(0.00)
$\log TV_{t-j}$			4.09***		59.53***	4.36***	
			(0.00)		(0.00)	(0.00)	

Note: Outcomes of the 3 autoregressive models, with a 6-day order: model 1 represents the dynamics between volatility and search volume; model 2 between search and trading volume; and model 3 is a joint model of volatility, search volume and trading volume. The models are described in Equation 3.4: $\mathbf{x}_t = \mathbf{c} + \sum_{j=1}^{p} \beta_j \mathbf{x}_{t-j} + \boldsymbol{\epsilon}_t$. The vector $\mathbf{x}_t$ accounts for volatility (σ_t^2), Google search volume (GSV_t) and trading volume (TV_t) time series, whereas $\mathbf{c}$ and $\boldsymbol{\epsilon}_t$ are respectively vectors of constants and shocks. All series are in their logarithmic form. The daily conditional volatility of the Ibovespa log-returns was estimated using a t-distributed GARCH(2,1) model, from 2004 to the 1Q16. Model coefficients are presented in Panel A and the statistics of the Granger causality $\chi^2(4)$ test in Panel B. P-values are in parenthesis. *, ** , *** denote significance at 10%, 5%, 1%.

responses of volatility and searches to their own shocks, as expected. In the chart (c), the response with slow decay is an evidence of the also persistent impact of a shock in search volume to volatility, which corroborates with the hypothesis of price pressure induced by attention. The chart (d) indicates that there is also an elevation in attention in response to a volatility shock.

3.4 Final Considerations of the Chapter

In this chapter, we investigated the effects of the attention of individual investors on stock market volatility. Attention is a requisite to sentiment and exacerbated sentiment can lead to a suboptimal allocation of the attention, which, in turn, results in behavioral biases (Kahneman, 1973).

Individual investors have relevant participation in the stock market. Our general conjecture is that an increased attention by this group will affect the market due to their fundamental nature and behavior. More noise trading shall result in abnormal short-term volatility, what Barber and Odean (2008) called an attention-induced price pressure.

We used as a proxy for the attention level of noise traders the amount of search queries performed at Google, a popular search engine. The capacity of this measure for this purpose was evidenced in studies from Dimpfl and Jank (2016) and Da et al. (2011).

We primarily adopted the Ibovespa for the research due to its popularity and representativeness of the Brazilian stock market. Besides, searches about stock indices tend to be less ambiguous than the ones for individual companies. We opt for the keyword “Ibovespa”, among similar ones, due to its relevance and better capacity of representing searches specifically for the index. Moreover, similar keywords presented search volume highly correlated with “Ibovespa”.

Our analysis resulted in considerable outcomes. After modeling the daily volatility of the Ibovespa, we verified whether lagged Google search volume determines this volatility, controlling for trading volume and serial correlations.

These results indicated that, indeed, there is a temporary volatility induction in moments preceded by higher retail attention. We found that Google search volume for the previous day explains current volatility in Brazil. In the US, the same evidence was found by Dimpfl and Jank (2016), despite relevant differences between the two countries regarding the degree of asset concentration and of individual investor influence.

We also found evidence that more trading volume arouses the attention of individual investors and determines higher index variance. In line with this finding, Barber and Odean (2008) and Da et al. (2011) proposed the use of search volume as measure of the attention level of investors. On the other side, Brooks (1998) and Dimpfl and Jank (2016) did not find such relationship in the markets where they address their analysis.

The findings corroborate with our research hypothesis: in moments of higher atten-

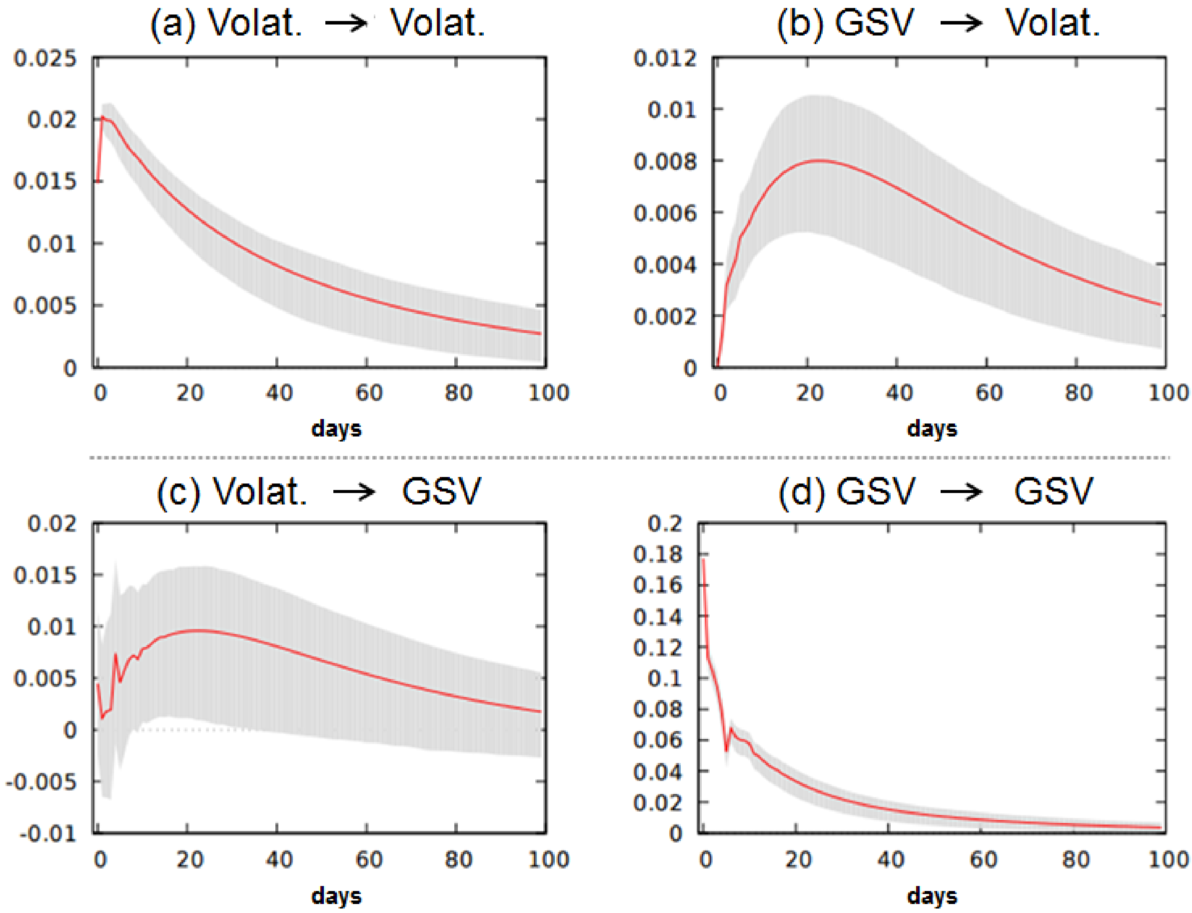

Figure 3.3: Impulse response functions of the series

Note: (a) Response of the Ibovespa conditional volatility (t-distributed GARCH(2,1) model using log-returns) to a shock in volatility. (b) Response of the volatility to a shock in search queries (Google Trends index). (c) Response of the search volume to a shock in volatility. (d) Response of the search volume to a shock in search volume. The results were obtained through a $VAR(6)$ model, using daily logarithmic series from 2004 to the first quarter of 2016. Shaded areas highlight a confidence interval of 95%. The impulse response was simulated for 100 days.

tion, the behavior of individual investors contributes to a volatility increase in the Brazilian stock market. As in the multiagent model described by Lux and Marchesi (1999), a more intense participation of noise traders alters the interaction dynamics among agents and transmits more instability to prices . Therefore, measuring the attention level of this influential group may contribute to a better understanding of the stock market.

Complementary studies may address the predictive capacity of the attention level on the activity of other markets, such as the money, Foreign Exchange (ForEx) or even labor ones. Another interesting approach would be the analysis of the attention level measure by the access or publications on popular social media, such as Twitter or LinkedIn. These emergent platforms have shifted the relations between financial services providers and users (Eldridge, 2016; KPMG, 2012). Promising results could also come from an investigation of the relationship between Internet search queries and analysts (or brokerage firms, in general) coverage or releases in the Brazilian context. The study of how attention affects abnormal short-term returns, especially in IPO events, may also result in interesting findings.

In addition, one can conceive trading algorithms that combine machine learning techniques with the diversity of robust databases that have been deployed across large platforms. These databases comprise data on users, documents, news, assets, as well as macroeconomic statistics. Propositions related to artificial intelligence and big data may result in investment techniques with better balance between risk and expected return.

References

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *The Review of Financial Studies*, *21*(2), 785–818 (cit. on pp. 10, 20, 25, 28, 43, 47, 50, 51, 68, 86, 98).

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73–92 (cit. on p. 48).

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, *45*, 150–156 (cit. on p. 53).

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528–543 (cit. on p. 48).

BMFBovespa. (2017). Participação dos investidores. Available at bmfbovespa.com.br. (Cit. on p. 60).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327 (cit. on pp. 60, 90).

Brooks, C. (1998). Predicting stock index volatility: Can market volume help? *Journal of Forecasting*, *17*(1), 59–80 (cit. on pp. 47, 68).

Burton, E., & Shah, S. (2013). *Behavioral finance: Understanding the social, cognitive, and economic debates*. John Wiley & Sons. (Cit. on pp. 10, 48).

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499 (cit. on pp. 10, 11, 20, 43, 47, 48, 52, 53, 56, 57, 60, 65, 66, 68, 75, 87, 97).

Damien, C., Ahmed, B. H. A. Et al. (2013). *Predicting financial markets with google trends and not so random keywords* (tech. rep.). (Cit. on p. 47).

De Souza, H. E., Barbedo, C. H. D. S., & Araújo, G. S. (2017). Does investor attention affect trading volume in the brazilian stock market? *Research in International Business and Finance* (cit. on p. 56).

de Ávila Marques, T., Machado Guimarães, T., & Macial Peixoto, F. (2015). A concentração acionária no brasil: Análise dos impactos no desempenho, valor e risco das empresas. *RAM. Revista de Administração Mackenzie*, *16*(4) (cit. on p. 58).

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, *22*(2), 171–192 (cit. on pp. 10, 47, 48, 54, 58, 59, 65, 66, 68, 75, 86–88, 97, 98).

Eldridge, R. (2016). How social media is shaping financial services. (Cit. on p. 70).

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007 (cit. on p. 60).

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, *61*(4), 2017–2046 (cit. on p. 49).

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393–408 (cit. on pp. 19, 21, 22, 51).

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, *56*(4), 1533–1597 (cit. on pp. 9, 48, 49).

Huddart, S., Lang, M., & Yetman, M. H. (2009). Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science*, *55*(1), 16–31 (cit. on pp. 10, 47, 51).

Johansson, E. (2014). Creating daily search volume data from weekly and daily data (cit. on pp. 31, 60).

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ. (Cit. on pp. 9, 68).

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part i* (pp. 99–127). World Scientific. (Cit. on p. 48).

KPMG. (2012). Social media in financial services. Available at kpmg.de. (Cit. on p. 70).

Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, *397*(6719), 498 (cit. on pp. 47, 49, 50, 59, 70, 86, 88).

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483–510 (cit. on pp. 19, 22, 23, 51).

Miragaya, F. (2017). O efeito dos retornos anormais nas buscas por informação dos investidores. *Masters Thesis* (cit. on pp. 56, 57).

Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: Insights from search engine query data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *368*(1933), 5707–5719 (cit. on p. 47).

Rodolfo, E. R., Barbedo, C. H. d. S., & Val, F. d. F. (2017). Relevance of the attention effect in the brazilian stock market. *Latin American Business Review*, *18*(2), 165–187 (cit. on pp. 56, 57).

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in asia-pacific stock markets. *Pacific-Basin Finance Journal*, *38*, 107–124 (cit. on pp. 29, 30, 32, 43, 51, 57, 58, 98).

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, *41*, 17–35 (cit. on pp. 28, 29, 32, 43, 51).

4 Ostrich Behavior: Effects of Investor Attention on the Volatility Asymmetry

> The true art of memory is the art of attention.
>
> Samuel Johnson (1709-1784)
> Writer, literary critic and poet

Volatility fluctuations are very notable in capital markets. The existence of clusters, for instance, has implications on pricing, risk management and market efficiency tests, being considered in several studies and models.

These fluctuations bring significant effects in required returns, which adjust current prices. They help to explain stylized facts, such as more common negative than positive returns (return asymmetry), more frequency of extreme returns than would be expected in a normal distribution, and higher volatility after falling than after rising prices.

This last evidence, called volatility asymmetry, was firstly addressed by Black (1976). While upward trends are typically more gradual, downward ones are notably steeper. The author argued that this asymmetry is a result of a higher leverage that occurs when prices fall. When the firm value decreases, the equity becomes riskier, which increases volatility. Schwert (1989) argued that the operational leverage also increases the intensity of this negative relationship between returns and volatility in bad times. Besides, an increase in stock trading leverage may lead to margin calls and forced selling, pushing prices down.

Other authors, however, verified that the magnitude of this effect is too limited to explain this negative correlation between current returns and future volatility. A second hypothesis was raised (or emphasized) by Pindyck (1984), French et al. (1987) and Campbell and Hentschel (1991). According to them, this anticipated increase in the volatility raises the expected equity return, causing an immediate decline in its price. In other words, they argued that the effect of prices coming from changes in volatility is more expressive than the impact in the opposite direction.

Campbell and Hentschel (1991) described how the volatility asymmetry can be explained by a feedback effect, based on future dividend shocks. Since positive shocks tend to be accompanied by other positive shocks, the first one generates an expectation of increase in volatility, driving an increase in the expected return, hence reducing the stock price, mitigating the positive impact of the price. If there is a negative shock, however, the price also decreases due to an expectation of an increase in volatility, but in this case this effect amplifies the negative impact of the shock. This amplification generates an excess of kurtosis and then extremely negative returns become more common than extremely positive ones. The authors represented this feedback effect using a variance model that helps to explain these asymmetry and kurtosis patterns of daily and monthly returns of US stocks over 63 years.

This hypothesis that the volatility asymmetry occurs due to fluctuations in expected returns takes as assumption a positive correlation between expected return and volatility, However, an opposite relationship was identified by Breen et al. (1989). Besides, it is reasonable to consider that, among volatility asymmetry determinants, there are market (systematic) factors as well as specific (idiosyncratic) factors of individual assets.

Identifying the determinants of volatility asymmetry is even more controversial in daily than in lower frequencies. Avramov et al. (2006) showed that the leverage effect might occur only in lower frequencies, since daily changes in leverage are transient and of smaller magnitude. Also, deviations in expected returns due to economic cycles are barely noticeable in daily series, when returns tend to be more unpredictable (Cochrane, 2001; Lehmann, 1990; Sims, 1984).

In this context, Avramov et al. (2006) studied the impact of trading operations on the daily volatility of stock prices. The authors verified that the activity of contrarian investors (who the authors considered as informed ones) reduces the volatility after a decrease in prices. On the other hand, herd behavior among investors, which results in less informed and liquidity-driven trading, increases volatility in this situation. This robust effect found in the relationship between volatility and lagged returns indicates that the two classical hypotheses might not be enough to explain the phenomenon of volatility asymmetry in daily returns.

Besides herd behavior, other behavioral aspects related to finance might be associated with daily volatility asymmetry. There is this hypothesis that some types of behavior create a feedback that induces decisions such as panic selling. Loss aversion (prefer avoiding losses to conquering equivalent gains) and endowment effect (tendency to hold losing stocks for a long time and sell winning stocks too soon) are aspects that generate emotionally charged attitudes, illogical from a financial point of view and that may bias risk and probability assessments (Tversky & Kahneman, 1979).

It is known that the attention of investors fluctuates over time (Da et al., 2011) and by itself determines an increase in volatility (Dimpfl & Jank, 2016). Would the intensity of investor attention have the power to accentuate or mitigate the perceived difference between volatility levels over good and bad times? This is the question we want to answer in this chapter.

Few studies addressed the impact of attention on volatility asymmetry. In a cross-country investigation, Talpsepp and Rieger (2010) verified that economic development and market efficiency reduce volatility asymmetry, while analyst coverage has a positive influence. Dzieliński et al. (2018) did a similar research, using a large sample of monthly returns, from 1989 to 2007, of US stocks. The number of analysts following a specific firm was adopted as attention measure. The asymmetry parameter was obtained through an Asymmetric Power Autoregressive Conditional Heteroskedasticity (APARCH) model. The results of the cross-section analysis showed that stocks with higher analyst coverage (and with larger dispersion among the forecasts) presented higher asymmetry. These results were expressive for stocks with low share of institutional investors and high idiosyncratic volatility. The leverage effect, documented by seminal papers, was not significant.

The authors associated this finding with an attention asymmetry that would be supposedly in the same direction of the volatility asymmetry. In other words, they assumed that investors become more attentive in bad times. According to them, this can be verified by the negative correlation between Google searches and returns, and by the surprising finding that hospitals attend more patients during bear markets (Engelberg & Parsons, 2016). Besides, due to the relationship between attention and volatility described by Andrei and Hasler (2015), attention is even more asymmetrical for companies that receive a higher level of it.

Recently, analysis based on information released over the Internet contributed to the identification of two different behavioral patterns that may explain how attention reacts to positive and negative news. One of these patterns, evidenced by Karlsson et al. (2009), is that individual investors tend to login less on their online accounts during bear markets. Kaustia and Knüpfer (2012), moreover, states that people are reluctant to share the results of bad investments with others (although it does not mean that they are less attentive). This phenomenon was named ostrich effect, an allusion to the legend that ostriches hide their heads in a hole when they are afraid. Gherzi et al. (2014), on the other hand, attest that the volume of logins raises both in good and bad times, suggesting that the investor behaves in a more vigilant way, analogous to that of a meerkat.

In this context, and trying to solve this controversy, the goal of this study is to investigate the longitudinal impact that fluctuations in the attention levels of investors induce in the volatility asymmetry of stocks. Our thesis is that more attentive investors reduce the difference in market volatility levels in good and bad times. When the market is bearish, volatility tends to increase, but a lower level of attention mitigates this effect.

We used for the analysis the returns of the most relevant index of the Brazilian stock exchange, since the asymmetry found in indices is usually higher than the ones of individual stocks (Andersen et al., 2001; Tauchen et al., 1996). Moreover, we performed the analysis in daily frequency since it is less vulnerable to economic cycles and leverage effects. Last but not least, we adopted as attention measure an aggregate variable — the daily amount of search queries performed at Google, a very popular search among non-sophisticated investors.

Our results confirm that more attention significantly reduces the asymmetry level of the market volatility, measured by an APARCH model. This outcome persists in the presence of several control variables and in different specifications of the volatility and regression models. The pattern also holds when we remove from this attention measure the correlation with professional attention (measure by the user activity level at Bloomberg, a sophisticated financial information provider), aiming to analyze only the non-professional attention dynamics.

Our approach contributes to a better understanding of the drivers of volatility asymmetry, a notable phenomenon in the stock market. In this sense, our purpose is to address the problem of the controversy of the identification of these drivers. Besides, it helps to analyze the fluctuations in the attention of investors and their effects on asset prices. Certainly, there is no consensus on whether attention is a cyclical or countercyclical variable, and our investigation helps to reach a conclusion.

Understanding the determinants of asymmetry levels and the implications of the dynamics of the agents' attention is useful, from the perspective of investors, in asset pricing and risk management. From the side of corporations, it aids in managing information releases and planning public offers. The relevance of this study is justified because our findings have the potential to increase the effectiveness of those processes.

In the next section, we present the most relevant references related to volatility asymmetry and to the dynamics between attention and asset prices. After that, we describe the research hypotheses, the sample, variables of interest and modeling decisions regarding both volatility asymmetry and the relationship between attention and asymmetry. The empirical results are then presented and discussed. In the end, we make final considerations.

4.1 Literature Review

In this section, we present one of the most relevant references used as theoretical and empirical background of this study. Firstly, we analyze papers that focus on volatility asymmetry and its determinants. After that, we expose studies and conjectures about the influence of the attention on the dynamics of asset prices.

4.1.1 Volatility Asymmetry

Due to the relevance that volatility asymmetry has in markets, some studies tried to investigate the major factors that determine it. One of those that gained higher prominence was Bekaert and Wu (2000), which developed a framework and an empirical approach to examine the asymmetry both at firm and market levels.

As major hypotheses for the asymmetry, the study mentions the leverage effect, presented and explained in the seminal papers of Black (1976) and Christie (1982). It shows that the increase in volatility when markets are bearish arises from an increment in the risk due to a decrease in equity values, and consequently a growth in firms leverage. Christie (1982) and Schwert (1989) evidenced this effect but recognize that it cannot be the only determinant for such high asymmetry patterns.

The volatility feedback is also considered, supported by the works of Pindyck (1984), French et al. (1987) and Campbell and Hentschel (1991). This phenomenon was described based on the idea that deviations on risk premiums naturally have an impact on the asymmetric profile of the volatility. The causality would be, in this case, in the opposite direction compared to the one of the leverage effect: an anticipated increase in volatility raises the expected returns on equity, when prices fall.

These patterns are supported by the idea that volatility is persistent, so shocks (both positive and negative) increase future and current volatility. Besides, it is based in an intertemporal relationship between expected return and conditional variance.

Considering these dynamics, French et al. (1987) and Campbell and Hentschel (1991) found a direct relationship between volatility and expected return. However, Turner et al. (1989), Glosten et al. (1993) and Nelson (1991) detected a correlation in the opposite direction. Other studies found a non-significant relationship. Besides that, according to what the CAPM determines, a condition for this hypothesis to hold at the firm level, it would be necessary that the market portfolio covariance respond positively to volatility increase.

The most important contribution of Bekaert and Wu (2000) is the test of these relevant hypotheses at the market and firm levels. The authors pointed out that the previous studies generally test the leverage effect at the firm (or portfolio) level, while the volatility feedback is tested using aggregate market data. Besides, Bekaert and Wu (2000) evidenced the influence of covariance asymmetry on the volatility asymmetry. When the covariance between market and stock returns increases in bad times, the feedback effect gets stronger.

The authors assume in the model that conditional volatility is persistent and that the conditional version of the CAPM holds. This means that the return excess of the market portfolio is the product of the price of risk and the market conditional variance, and the the stock return excess of any firm is the price of risk multiplied by the conditional covariance between firm and market return.

The model considers the (simultaneous) effects of leverage and volatility feedback, among the mechanisms that induce asymmetry, coming both at market and firm levels. Therefore, if bad news at market level appears, two effects take place. Firstly, while news are an evidence of higher market volatility, investors will probably also revise conditional volatility given its persistence. This upward revision of the market variance has to be compensated by a higher expected return, leading to reductions on prices and market values. This negative shock in the return generates an increase in the conditional variance. Besides that, it results in a higher general market leverage, and, consequently, in higher volatility. This means that in this case the leverage effect increases the volatility feedback effect.

At the same time, the resulting impact arising from the release of good news is not so clear. In this situation, there will be an increase in the current volatility and an upward revision in conditional volatility. This increment requires a higher expected return, resulting in a reduction in prices, which can cancel the initial positive shock. Hence, in this case, the feedback effect diminishes the initial effect of the volatility. Besides that, this positive shock elevates prices, reducing the general leverage and the conditional variance at the market level.

At the firm level, these dynamics of the initial impact of news is basically the same. However, the volatility feedback effect shows differences. The existence of this effect depends on an increase in the covariance between the stock and the market returns, in response to market shocks. If the shock is totally idiosyncratic, only the leverage effect generates asymmetry, because the covariance does not change, neither does the expected risk premium.

This impact on the unconditional covariance typically appears among firms. The higher the firm systemic risk, the higher should be the increment in the conditional covariance of its stocks due to market shocks. This leads to an increase in the required return, completing the feedback cycle of the volatility, which is also positively influenced by the firm leverage ratio.

The most relevant proposition of Bekaert and Wu (2000) is that the covariance asymmetry accentuates the volatility feedback effect. The authors argue that the covariance asymmetry was not properly investigated by previous studies. The proposed model, therefore, specifies this asymmetry arising both from the leverage and volatility feedback effects, given their impact on the variance asymmetry at the market and firm levels. The beta asymmetry is also considered in their framework, although this circumstance is less salient (generated by idiosyncratic shocks, but not so by systemic shocks) and is rarely treated in the models.

The authors adopted the conditional CAPM to investigate the interactions between expected values and variances of the stocks and the market, and built portfolios grouping firms with similar leverage ratios. The price of risk was defined according to the CAPM at the firm (and not the equity) level. The CAPM parameters were defined using a multivariate GARCH model, specifying a variance-covariance matrix from a asymmetric version of the Baba-Engle-Kraft-Kroner (BEKK) model. The model at the firm level allows a clear segregation of the leverage and the volatility feedback impacts. This setup results in a large number of parameters, but some constraints significantly reduce this amount.

This specification leads to variances and covariances exactly the way the Christie (1982) model describes with riskless debt in the case with constant firm variances. Fluctuations in those variances impact the deviations in the stock volatility when leverage is also higher.

At the market level, the volatility follows an univariate asymmetric GARCH model, adjusted by leverage. In addition to the same mechanism of Christie (1982) model, the influence of the leverage ratio in the conditional variance model occurs in two ways: through shocks of similar firms, which generate volatility effects when leverage increases, and through a leverage growth in the previous moment, which elevates the GARCH effect.

At the portfolio level, volatility unfolds in three components: one that adjusts the ARCH factor upwards only when the current leverage of the portfolio is higher than the previous one of the market; and other two ones that involve past idiosyncratic covariance and variance that adjust similarly.

The generalized BEKK model accounts for covariance dynamics with: a constant term that represents the leverage effects of Christie (1982); an autoregressive variance term, influenced by leverage; a term that represents the covariance persistence; and shock components, whose effects depend on the combination of individual and market shocks. The generalization imposes non-linear restrictions in the parameters and, consequently, in the particular magnitude of the responses, as well as implies that variances and covariances are defined by the same parameters.

After verifying that the model is well specified, the authors obtained for the empirical approach daily data from 1985 to 1994 of the prices and market capitalizations of 172 firms that comprise the Nikkei 225 index, as well as biannual data of the book value of their debt.

Three portfolios with five stocks each were built according to the leverage, excluding financial institutions. Despite measurement errors, the portfolios show very distinct proportions over the period. The one-month Gensaki was used as short-term interest rate.

The authors used likelihood ratio tests in order to verify the potential validity of models that are more restrictive than that one that simultaneously considers the presence of leverage factors, asymmetric shocks and volatility persistence. The results indicated that the leverage measure is not determined solely by the behavior of the volatility of the Japanese stock returns. Even removing leverage effects, the asymmetric volatility holds.

The volatility feedback can be generated by the dependency that firm covariance and volatility have over market shocks. The results suggested that the asymmetry effects are wider than simply feedback dynamics, or merely indicate correlation between market and firm shocks.

The parameter estimates of the model indicated a relevant persistence in the conditional volatility, at the market and portfolio levels. At the market level, return shocks show expressive effect over volatility asymmetry. The asymmetry is caused substantially by portfolio shocks in the low leverage portfolio, though this asymmetry is not very significant. Medium and large leverage portfolios exhibit an asymmetry of higher magnitude caused by market shocks. Impact curves indicate that the leverage effect accentuates the asymmetry, but the influence is secondary when compared to the feedback effect.

As stated by the time-varying risk premium theory, the conditional covariance has an important role in determining the expected excess return and the volatility feedback. Therefore, the authors verified whether negative shocks at the market level lead to an increase in the covariance between the market and the portfolios, particularly the medium and high leverage ones.

In general, the results showed a persistence in the covariances. More important, the high leverage portfolio showed an elevated covariance asymmetry. The high leverage portfolio covariances increase only when the market and portfolio shocks are of the same sign (and increase substantially when both are negative), while, in the medium leverage portfolio, they increase only when the portfolio shock is positive and the market one is negative.

The authors evidenced that the volatility asymmetry of high and medium leverage portfolios are in fact related to the asymmetric response of the covariance due to market shocks, and these effects are elevated by leverage.

The analysis of the beta responses to the shocks indicates similar patterns only in the medium and high leverage portfolios. Portfolio and market shocks of the same sign increase betas, but different patterns were found in the low leverage portfolio. The results suggest a leverage effect in this portfolio beta, but this occurs solely due to a lack of a relevant volatility feedback effect. In general, the authors conclude that the feedback effect is the one that determines the beta dynamics.

The economic significance of the variance asymmetry was assessed by analyzing the effect of the shocks on the series average values. Regarding the volatility, portfolio shocks generate strong asymmetry in the low leverage portfolio, but total volatility asymmetry remains low. Still, the difference between the effects of positive and negative combined shocks is 45 basis points. In the high and medium leverage portfolios, the difference is 96 and 153 basis points, respectively.

With respect to covariances, the effect is clearer because portfolio shocks typically generate an expressive asymmetry in all the portfolios. Adopting a measure of unconditional price of risk, the risk premiums range from 12 to 55 basis points for market and portfolio shocks combined, which is relevant given the average return excess of 1.73% per year in the sample of Japanese firms.

The volatility feedback induces return shocks high enough to compensate the new expected return, which is even higher for negative shocks after normalizing shocks of different signs. In percentage points, the difference is still more visible and is irrespective of the price of risk. An increase in volatility raises by 16% the expected return due to bad news and by 5% due to good news. When the higher level of uncertainty is priced and there is an increase in the covariances, combined negative shocks generate an increase of 17% in the risk premiums, while positive shocks increase only from 5% to 8%. With respect to betas, the simulations did not find evidence of asymmetry, except for lower magnitudes in the high and medium leverage portfolios.

The analysis using impact curves shows similar results. For equivalent shocks, the low leverage portfolio presents a volatility asymmetry much higher. This occurs because its firm shocks are much higher than the ones of the other portfolios. The fact that the shocks maintain high asymmetry corroborates with the reasoning that feedback effect dominates the leverage one.

An additional test was performed by the authors to verify whether size influences the results. They divide the sample in three groups based on market capitalization. The stock portfolios of higher and lower value were then subdivided according to the leverage ratio. The test steps were similar to the previous analysis. Specification tests did not reject the new model and the conclusions remained, showing that a strong volatility asymmetry is still present after removing the leverage effect. The effects found were economically significant, and the covariance asymmetry presented a direct and relevant influence in the risk premiums among all the portfolios. The beta asymmetry was not identified or was too weak. Besides that, they did not find evidence that confirm the findings of Cheung and Ng (1992) in the US market, that the volatility asymmetry is higher for stocks of small firms.

The findings of Bekaert and Wu (2000) were striking because they evidenced the feedback effect in the variance asymmetry and refuted the hypothesis that the leverage effect is preponderant for this phenomenon. However, the authors made clear that other factors should also determine this asymmetry. Our work intends to verify the contribution of investor attention for these dynamics. In the next section, previous studies with connected approaches are presented.

4.1.2 Attention Reaction to the Asset Prices

The model of Andrei and Hasler (2015) associates attention with market variables (returns, volatility and risk premium) to understand its role in determining prices. Besides evidencing that investor attention is very sensitive to recent experiences, the parameter estimates indicated that attention is high in bad times. They interpreted this finding supposing that investors do not have incentives to make efforts to learning during an expansionary economy. During a recession, the perspective of a reduction in future consumption captures more investor attention to estimate more accurately changes in fundamentals.

However, the authors recognize that this evidence is not conclusive when the results of other studies are taken into account. While Patton and Timmermann (2008), Da et al. (2014) and Garcia (2013) indicate that forecasts are more accurate during crises, the findings of Van Nieuwerburgh and Veldkamp (2006) indicated the opposite. Besides that, there are periods of economic expansion, such as the end of the 1990s, when the high level of media coverage suggests exalted investor attention.

Another important evidence of attention cyclicality is the model of Karlsson et al. (2009), which describes interactions between stock market variables and the attention level of investors. Their novel approach allows connecting an observable behavior (the decision about obtaining information) to internal psychological variables. These variables are not observable, but are of great importance in the fluctuation of the investors' preferences in good and bad times. The model extrapolates the investment environment, fitting any situation in which people care about information but have some ability to protect from them.

As in previous studies (Backus et al., 2004; Barberis et al., 2001), they propose a model that incorporates psychological aspects in price variations. However, a new data source is used, corroborating with the idea that people derive their utility directly from information about wealth changes. The idea is to investigate how investors carried with emotions and limited in the assimilation capacity process released information.

In the model, a sole investor is represented with some level of control between the timing to access specific information about their wealth and the effect of this information in their utility. The investor correctly interprets any information he accesses and accurately evaluates the impact of potential information in this sentiment.

In other words, the investor decides whether he awakens his attention or not to obtain more precise information about the position of his investments, conditioned to the general previous market news that he naturally receives. This awakening contemplates both the psychological processes and the necessary behavior for that.

The authors define two effects of selective attention on utility. The first one, named impact effect, corresponds to an increase in the psychological impact of the information on the utility. This effect is based on the prospect theory, which determines that the utility depends on how the results deviate from a pre-specified reference point. Among other factors already documented in previous studies, they argue that attention amplifies the marginal impact, both of losses and gains, in the utility.

The second effect of the attention is on the reference point update. The more attentive, the faster the investor updates this benchmark. This effect is supported by previous studies that indicated, for instance, that the reference points are more responsive to deterministic than to probabilistic information. Accessing more precise and specific information about this wealth would have, according to the authors, similar impact.

With this, the decision making model developed by the authors presents two moments of time. In the first one, there is a shock in the investor wealth. This shock consists of a component about which the investor learns automatically (having, good, bad or neutral content), and another one about which he can decide to learn (with no cost) or not. This discretionary component can assume a good or bad state with respect to the automatic component. At time 2, there is an additional shock in the final wealth of the investor, which can be a good or bad change in relation to the wealth at the first moment.

Deciding not to learn about the discretionary component at time 1 means burying the head under the ground (hence the name ostrich effect) and waiting to be aware of the content of this component only in the second moment. If he does that, his perceived wealth may not be equal to the actual position of his net worth.

The model relies on some assumptions that simplify reality. For instance, in the second moment, when the investor is psychologically attentive, his perceived wealth is always equal to the actual one. Besides that, the investor is risk averse in these preferences with respect to information about his wealth, and his utility at each time is centered at the level of the exact previous moment. The utility is then disturbed by the deviation of this perceived wealth compared to a previously determined reference point.

The basic premise of the model is that the investor conditions his decision of when to learn about the discretionary component to what he learns automatically about his wealth. If the automatic component has a good content, this decision is a trade-off between the advantages of receiving expected good news and the advantages of a more slow update of the reference point at the second moment, reducing the chances of disappointments. In this case, the investor will be attentive if the utility difference (between being and not being attentive) is sufficiently large and if the benchmark revision is significant enough. If the impact effect is equivalent to the effect of a slower update of the reference point, the investor chooses to be psychologically attentive if his degree of risk aversion is not very high.

Another possible scenario is the natural absorption of information of neutral content. In this case, comparing the value functions of an attentive and an inattentive investor , it is always more advantageous being inattentive in the first moment, due to his loss aversion. The change in expected utility at the second time will be identical, regardless of the investor being attentive or not in the first moment.

Lastly, when the automatic information has a negative content, the impact effect favors being inattentive while the updating effect of the benchmark justifies being attentive (in order to have a lower benchmark in the second instant). Hence, the investor will not be attentive if the impact effect is higher enough and if the reference update when inattentive does not take too long. If both effects are equivalent, the optimal solution is being always attentive to the discretionary component in this type of situation.

Regardless of that, the authors recognize that there is an indirect demand for information that serves as fundamental input for the buy and sell decisions by the investors. They take into account the expected utility of this demand to analyze the combined expected utility of checking the portfolio in the first instant. In other words, checking the value of the personal portfolio in bad times has psychological costs that makes it less probable, but can be justifiable in some cases, especially due to the heterogeneity of the investors.

Improvements in the model include a higher number of instants, time discounting the wealth value and relaxing the assumption that the investor is always attentive at the second moment. The authors argue that these adjustments to the model, in general, would increase the benefit of not being attentive to the information in bad times compared to regular times, corroborating even more with the ostrich effect.

Selective attention is, according to the authors, a rational mechanism given that investors are psychologically affected by assessed information. Schneider (2001) claims that less accurate information is perceived as less salient or vivid, having more room for self-manipulation of expectations with respect to knowledge. In other words, the authors consider that there are multiple ways to experiment with information.

The authors analyzed three different samples of Scandinavian countries to verify whether empirical data confirm the model outcomes: one with the amount of checkings on online accounts by investors from October 2003 to January 2004 in a large financial services company in Norway; another one with the quantity of investors logins on the funds investment position section in a large Swedish bank from June to October 2003; and a last one with the volume of checkings on the pages that inform the personal investments position in pension funds, at the pension Swedish authority, from January 2002 to October 2004.

The authors regressed the number of logins on the value of relevant stock indices (the current position and the average value of the last six days), as well as variables to control for alternative explanations. Deviations on the stock indices values were considered as a proxy for public (or automatically perceived) news.

The regressions with different databases indicate statistically and economically significant values for the coefficients that indicate the relationship between the indices values and the number of accesses. This pattern remains with subdivisions of the sample and the inclusion for day of the week and other variables to indicate logins that are merely to pay bills or accesses solely to the main page, as well as to indicate the number of transactions (removing the effect of an indirect information demand). The amount of accesses after bad news were higher than after news with neutral content only in the mutual funds database. The results were strongly in line with the ostrich effect suggested in the model.

Karlsson et al. (2009) argued that the results they found cannot be attributed solely to the fact that the investors are consuming the utility of good news. Based on evidence that results that do not reach expectations evoke disappointment (Bell, 1985; Gul, 1991; Loomes & Sugden, 1986; Zeelenberg et al., 2000), the authors argued that, after consuming good news about some wealth that will be realized in the future, people alter their expectations about this future wealth.

The media coverage asymmetry also may not totally explain the results, since the model compares the index value with its average in the week before and part of this asymmetry can be explained precisely by the ostrich effect in information consumers.

Another possible justification is the fact that investors are more prone to access the information providers when they want to buy stocks, and this aggregate effect would lead to an elevation in prices. The authors dismiss this argument when they show in one of the samples that the correlation between transactions and the index value (controlled by the number of accesses) is very weak compared to the correlation between accesses and the index value (controlled by the number of transactions).

The model of Karlsson et al. (2009) implies that the reference point of loss aversion should vary more (in module) in up than in down markets, leading to also asymmetry dynamics in the market risk premium. This effect helps to explain the pattern found by Griffin et al. (2004). They evidenced that positive returns lead to significant increases in the trading volume ten weeks after, in a relevant sample of countries. This notable phenomenon cannot be completely explained by liquidity effects, overconfidence and endowment effect. The drastic reduction in liquidity during crises are also supported by the ostrich effect, which determines that investors ignore the market in these moments to avoid having to mentally deal with painful losses. The positive asymmetry of attention can also be one of the reasons why there are subtle fluctuations in social transmissions of information, exacerbating crises in bad times or creating bubbles in good ones.

The ostrich effect makes people check their portfolio more frequently in good times, while they stay information averse when the market is bullish, taking the risk of losing an opportunity to optimally rebalance their portfolio or losing favorable chances to trade assets. From an evolutionary point of view, selective attention helps people to deal with risky investments, reducing short-term concerns and allowing long-term gains. This might reduce at some level the risk premiums and have effects that can counterbalance other biases.

This behavior pattern also introduces effects on the dynamics of volatility. We know that attention induces an increase in volatility, but the ostrich effect gives to attention a cyclical nature (lower in bad times). Since volatility has a notably countercyclical pattern (higher in bad times), we hypothesize that the cyclicality of attention mitigates

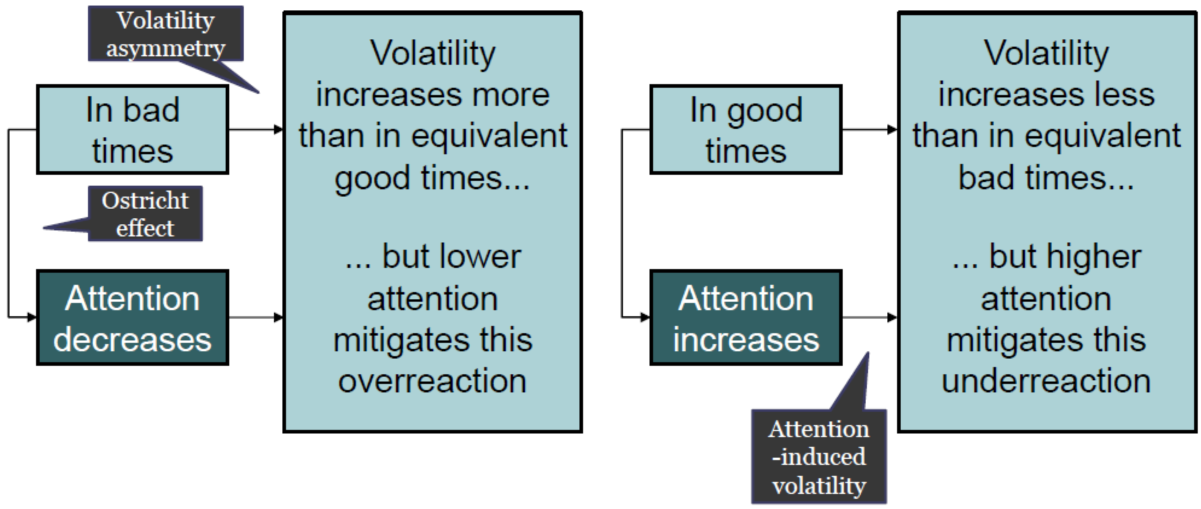

Figure 4.1: Illustrative model of the relationship between attention and volatility asymmetry

Note: Flow chart illustrating the logic that supports our conjecture that attention reduces volatility asymmetry. The model combines the ostrich effect, evidenced by Karlsson et al. (2009), the volatility asymmetry, demonstrated by Black (1976), and the attention-induced volatility hypothesis, presented by Dimpfl and Jank (2016) and Lux and Marchesi (1999).

volatility asymmetry.

In other words, when investors are more attentive, there is a lower imbalance between volatility levels that occur in economic movements of contraction and expansion. In bad times, volatility raises more than it does in equivalent good times. However, attention has a opposite behavior. Since attention and volatility are positively correlated (Dimpfl & Jank, 2016), this opposite behavior softens the relative excess in volatility increase.

In favorable periods, the opposite occurs. Volatility increases less and attention increases more than in equivalent unfavorable times. By increasing market volatility, attention mitigates the relative volatility underreaction in bad times. Figure 4.1 illustrates this expected impact of the ostrich effect on the volatility asymmetry.

This model subsidizes the main research hypothesis of this study:

H_1: An increase in attention over time leads to a reduction in volatility asymmetry

Since previous studies (Barber & Odean, 2008; Dimpfl & Jank, 2016) state that it is substantially the attention of non-professional investors that induces most of the non-fundamental volatility, we establish a second hypothesis, that the attention of this investor class is the one that fulfills the role of asymmetry mitigator:

H_2: Higher retail attention over time leads to a decrease in volatility asymmetry

We describe in the next section the methods performed and decisions made to test the hypotheses outlined, as well as the time series that are part of the sample.

4.2 Methodology

In this part of the study, we present the sets of secondary data obtained to perform the tests used to investigate the conjecture that the ostrich effect influences the fluctuation of the stock market variance. We describe the steps and specifications of the analysis, as well as the expected signs and variables of interest.

Specifically, our purpose is to evaluate whether there is an effect of investor attention on the daily asymmetry of the volatility, using the Ibovespa as a reference. In this context, the next section presents details on the return series obtained to estimate the volatility asymmetry, and the Internet search data used as proxies for measuring investor attention. We also describe the control variables that might have an influence on the volatility asymmetry.

4.2.1 Data Description

We used daily time series of the Bovespa Index (Ibovespa), very popular and that encompasses the most representative stocks of the Brazilian market. The series cover a period of 14 years, between 2005 and 2018, resulting in 3,459 days, a reasonable enough interval for this type of analysis. The series used to measure investor attention are not available before this period.

The volatility index was estimated with respect to the dividend-adjusted log-returns of the Ibovespa. As we detail in the next section, the daily series of the asymmetry index is one of the parameters of the volatility model, performed in moving windows. Given the necessity of an initial time window for this procedure, a part of the sample period is not regarded in the assessment of the relationship between attention and asymmetry.

Following several previous studies (Bank et al., 2011; Da et al., 2011; Dimpfl & Jank, 2016; Jacobs & Weber, 2012; Kita & Wang, 2012; Vlastakis & Markellos, 2012), the investor attention is measured by the volume of search queries performed at Google. Since Google is by far the most popular web search engine in Brazil, the index that represents the amount of searches made by users to finance related keywords is a proxy of the retail investor attention. More details regarding this variable are presented in Chapter 1.

We chose as keyword the term "bovespa" due to its larger representativeness. Bovespa is the previous name of the Brazilian stock exchange B3 until March 2017. Despite this rebranding, the name is still widely used, especially to refer to the main stock index, Índice Bovespa (Ibovespa), which curiously did not suffer a name change, due to the popularity.

As Google Trends does not allow extracting daily series covering such a long period in a single query, we obtained the data covering each quarter and concatenated the series using the normalization proposed by Chronopoulos et al. (2018). The index ranges from

0 to 100, but were converted to a logarithmic scale to be used in the regressions (results remain in the original scale). Three observations in the beginning of the sample with null values were discarded. The search volume and financial databases were combined excluding days when any of the information were not available, meaning that only trading days were considered.

Also following previous studies, as control variables, we considered typical risk factors (size, *momentum* and price-to-book ratio) and the leverage ratio. Size was measured by the log of the market value of the Ibovespa. *Momentum* is the aggregate return over the last twelve months, excluding the one to which the observation refers. The price-to-book is the ratio of market to book values of the index. Leverage is the ratio of aggregate debt to the equity value of the index. Variables that depend on book values update less frequent than a daily basis due to the availability of information. In this case, we repeated the observations until new information is available. Ibovespa variables are obtained by summing the values of the stocks that comprise it, weighting by the share of each stock. These financial information are in nominal values and were obtained from Bloomberg.

We also obtained an index that represents the amount of search for information about the most representative stocks at the platform Bloomberg. This series is a proxy of the professional investor attention, given that a representative part of Bloomberg users belong to this class (Ben-Rephael et al., 2017). Less sophisticated investors are more susceptible to behavioral biases (Feng & Seasholes, 2005) and have a preponderant role in the induction of non-fundamental volatility (Dimpfl & Jank, 2016; Foucault et al., 2011; Lux & Marchesi, 1999). Hence, as an additional test, we verified the effect on asymmetry both from professional and non-professional attention.

Using the variable "NEWS HEAT READ DMAX" of the Bloomberg platform, which indicated the number of times each article is read by the users, combined with the number of times the users actively performed searches for news for a specific stock. This index varies from 0 to 4. Missing values were replaced with 0. More information about this measure can be obtained in Chapter 1 and by Ben-Rephael et al. (2017).

Since there is no Ibovespa search volume at Bloomberg, we performed a Principal Component Analysis (PCA) of the series of the five most representative stocks of the index (measured by the theoretical quantity) on the composition in September 2019: B3 (the stock exchange), Bradesco, Itaú (banks), Petrobras (oil company) and Vale (mining company). We considered the most liquid Petrobras ticker, Petr4, disregarding the also representative ticker Petr3.

Table 4.1 presents the descriptive statistics of the series over the whole interval. The index presents an average return of 0.05%, with an standard deviation of 1.72% and a range of 26.05%. The average market cap of the index was close to R$ 70 billion (around USD 18 billion, considering the average exchange rate of 2019), but it varied from R$ 23 to 148 billion over the period. With respect to the other classical risk factors, the

Ibovespa had an average *momentum* of 11.69% and market value 57% higher than the book value. In terms of leverage, the debt represented on average around 122% of the equity, fluctuating between half and double that amount. The index that represents the search volume of the keyword "bovespa" performed using Google presented an average value of 14 and a standard deviation of 10.

Table 4.1: Descriptive statistics

	r_t %	$Size_t$ BRL Billion	$\log Size_t$	Mom_t %	$P2B_t$	$D2E_t$ %	GSV_t	$\log GSV_t$
Average	0.05	69.68	4.81	11.69	1.57	122.23	14.13	2.45
Median	0.07	63.31	4.80	10.23	1.53	107.23	10.00	2.30
Standard deviation	1.72	28.93	0.18	24.58	0.37	46.57	10.08	0.60
Minimum	−11.39	23.61	4.37	−42.72	0.89	56.01	4.00	1.39
Maximum	14.66	148.81	5.17	78.55	2.75	206.97	100.00	4.61
Asymmetry	0.14	0.77	−0.17	0.29	0.58	0.44	2.10	0.56
Kurtosis	5.76	0.02	−0.35	−0.17	−0.27	−1.26	7.38	−0.64

Note: Summary of the characteristics of the daily series relative to the Ibovespa, used in the study, considering the whole period, from 2005 to 2018. From left to right, we present log of nominal returns, adjusted by dividends; the market value of the index, aggregating the capitalization of the stocks that comprise it, weighting by the respective shares; the log of this market value; the index *momentum*, which is the cumulative return over the last 12 months, except the most recent one; the ratio of the market value to the book value of the index (price-to-book ratio); the ratio of the book value of debt to the market value of the equity (debt-to-equity ratio); the Google search volume of the keyword "bovespa"; the log of this volume. The book values of the stocks that comprise the index are aggregated, weighting by their representativeness, to obtain the book value of the index.

4.2.2 Variance Asymmetry Modeling

Using the studies from Dzieliński et al. (2018) and Talpsepp and Rieger (2010), we estimated the variance of the Ibovespa returns through an APARCH model, developed by Ding et al. (1993). The APARCH is a variance model that presents some stylized properties of financial time series. The unconditional distribution has excess kurtosis and the model features volatility clusters and long memory in the returns. As the name says, this model (as well as other GARCH variations) captures the volatility asymmetry, meaning that it assumes that the variance raises more than returns are negative, in comparison to when they are positive and equivalent in magnitude (Campbell & Hentschel, 1991).

The following equations describe a general autoregressive–moving-average model (Autoregressive–Moving-Average (ARMA)) for the expected returnsr_t, and a general APARCH model for the variance σ_t^2.

$$
\begin{aligned}
r_t = \mu + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t \\
a_t = \sigma_t \epsilon_t \\
\sigma_t^\delta = \alpha_0 + \sum_{i=1}^{s} \alpha_i (|a_{t-i}| - \gamma_i a_{t-1})^\delta + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^\delta
\end{aligned}
\tag{4.1}
$$

The model of the expected value of the returns is characterized by shocks a_t in each period t, with finite unconditional variance, and a constant term μ. The model order, or the maximum lag of the terms, is defined by q autoregressive components, $\phi_i r_{t-i}$, and q moving average components, $\theta_i a_{t-i}$. Besides the order, the model parameters are μ, ϕ_i and θ_i. The term ϵ_t is an i.i.d. variable with zero mean and unitary variance, representing the distribution of the errors. Our main analysis encompasses both a version without orders in the ARMA model and another one that uses the Akaike criteria to find the most appropriate orders.

The variance model, on the other hand, is composed of a positive constant term α_0, an also positive power term δ, and the components $\alpha_i(|a_{t-i}| - \gamma_i a_{t-1})^\delta$ and $\beta_i \sigma_{t-i}^\delta$. Besides the parameters s and m, which define the orders, the constant term, α_i, β_i (both non-negative), δ and γ_i characterize the model. The component γ_i represents the asymmetry level of the model, assuming values from -1 to 1. In the APARCH model, it is our variable of interest.

Depending on the values of δ and γ_i, the APARCH model reduces itself to more simplified models, such as the GARCH one, proposed by R. Engle (1982) and generalized by Bollerslev (1986) and Taylor (1986), the TS-GARCH one, from Taylor (1986) and Schwert (1990), the GJR-GARCH one, from Glosten et al. (1993), the T-GARCH one, from Zakoian (1994), the N-GARCH one, from Higgins and Bera (1992), and the Log-ARCH one, from Geweke (1986) and Pentula (1986) (Gasparini et al., 2013).

To obtain the time series of the asymmetry parameter, we estimated the APARCH model several times using moving windows of the return series. In each estimation outcome, we extracted the parameter γ. After some trials, we chose to adopt a window of 2 thousand days, which results in an adequate remaining sample size with the rate of convergent results in the volatility modeling. Before being submitted to the moving window modeling, the returns were winsorized at the levels of 1% and 99% to maximize the convergence rate.

4.2.3 Modeling the Relationship between Asymmetry and Attention

After obtaining series of the volatility asymmetry index (γ_t) of the Ibovespa, we proceeded with the regression using, as explaining variables, the investor attention index (Google Search Volume, GSV_t) and control variables, as presented in Equation 4.2.

$$\gamma_t = c + \psi \log GSV_t + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t \tag{4.2}$$

The control variables are size, *momentum*, price-to-book ratio and debt-to-equity ratio. Following previous studies, we expected that moments of lower market capitalization and higher leverage lead a higher variance asymmetry of the index. The same is expected of moments of lower *momentum* and higher price-to-book ratio. Given the temporal sensitiveness of the data, we included year fixed effects (η_t) and standard errors clustered by year.

The Durbin Watson test for panel data pointed autocorrelation in the model residuals (an estimate of 0.31, with a p-value under 1%). Hence, we also performed an estimation by the Generalized Least Squares (GLS) method, elaborated by Pinheiro and Bates (2000).

The explaining variable of interest is $\log GSV_t$. Given the effect of attention on volatility and the positive asymmetry of volatility, we expect that lower levels of attention lead to higher asymmetry indices over time, corroborating with the existence of the ostrich effect. Hence, we expect that the coefficient ψ have a negative and statistically significant sign.

Regarding the control variables, we start from the relationship found by Dzieliński et al. (2018): negative effects of size and *momentum* and a positive effect of the price-to-book ratio on the asymmetry level. With respect to the leverage ratio, despite the controversy, our reference is the positive impact on asymmetry described in the seminal studies of Black (1976) and Christie (1982).

4.3 Empirical Results

In this section, we describe the outcomes of modeling the return variance of the Ibovespa, followed by the results of the regression that assess the effect that fluctuations in the investor attention induces to the longitudinal asymmetry of this variance.

4.3.1 Variance Asymmetry Modeling Outcomes

Table 4.2 presents the parameters estimates of the Ibovespa variance model, developed using the daily returns, covering the whole sample. We present both the version without orders in the ARMA model, as adopted by Dzieliński et al. (2018), and a version with an ARMA(1,1) model. The Ljung-Box and Goodness-of-Fit tests outcomes suggest adequate specifications, as well as the (very close to each other) information criteria estimates. The results are robust to less parsimonious models.

The results show marginally significant coefficients, of similar magnitude and opposite sign, for the parameters of the ARMA(1,1) model. In the variance model, the constant term was the only one that did not present statistical significance at 1%. The main variable of interest, the γ, presented an estimate of 0.56 and 0.58. Positive values are in line with the phenomenon that is commonly described in previous studies, that the volatility is higher in bad times than it is in good times with equivalent magnitude.

The asymmetry series were built by performing each window of 2 thousand days to this modeling. This means that an APARCH(1,1) model was generated comprising r_1 to r_{2000}, another one from r_2 to r_{2001} and so forth, until covering the whole sample. The last moving window was from r_{1460} to r_{3459}. The returns were winsorized at 1% and 99% for the modeling process.

For each generated model, we extracted the asymmetry index (γ_t). Therefore, the modeling process was performed 1,460 times. In very few cases, the modeling process did not converge, or generated values so extreme (outside the interval from -0.98 to $+0.98$) that suggests a convergence error. Hence, those values were discarded. This process resulted in 1,452 observations of the γ_t in the APARCH(1,1)ARMA(0,0) model and 1,404 observations in the APARCH(1,1)ARMA(1,1) model.

Due to the high number of iterations required by the maximum likelihood estimations, this process required a heavy computational load and took from 30 to 90 minutes to cover the whole sample.

Figure 4.2 presents the frequency distribution of the series of asymmetry index, after adopting these parameters. The index presents an average value of 0.49, standard deviation of 0.15, Q1, Q2 and Q3 of 0.37, 0.46 and 0.61, respectively.

The statistical parameters of the series indicate a relevant fluctuation in the daily volatility asymmetry. The heterogeneity suggests an influence of different variables in these dynamics. The values differ from those found by Dzieliński et al. (2018) in monthly US stock data from 1989 to 2007. Although most of the values they found were positive, the data they obtained were closer to a normal distribution, with an average of 0.18 and standard deviation of 0.20. Around 10% of the sample presented negative values, until -0.32. Naturally, the analysis of stocks, instead of aggregate indices, and a sample that covers a longer period should result in a set of asymmetry indices with a higher range.

Table 4.2: Asymmetry Modeling of the Ibovespa

	APARCH(1,1)ARMA(1,1)	APARCH(1,1)ARMA(0,0)
ϕ_1	0.60^{*}	
θ_1	-0.61^{*}	
ω	0.00	0.00
α_1	0.06^{***}	0.06^{***}
β_1	0.92^{***}	0.92^{***}
γ	0.56^{***}	0.58^{***}
δ	1.50^{***}	1.49^{***}
Distrib.	11.10^{***}	11.10^{***}
Log likelihood	9,588.66	9,588.47
Ljung-Box Tests		
R		
Lag[1]	0.14	0.00
Lag[2*(p+q)+(p+q)-1][5/2]	1.78	0.21
Lag[4*(p+q)+(p+q)-1][9/5]	3.06	2.21
R^2		
Lag[1]	3.22	3.14
Lag[2*(p+q)+(p+q)-1][5]	5.08	4.95
Lag[4*(p+q)+(p+q)-1][9]	6.22	6.07
Lagrange Multiplier Test		
ARCH Lag[3]	0.13	0.12
ARCH Lag[5]	0.13	0.12
ARCH Lag[7]	1.10	1.08
Information Criteria		
AIC	−5.54	−5.54
BIC	−5.52	−5.53
SIC	−5.54	−5.54
HQIC	−5.53	−5.54
Pearson Goodness-of-Fit Test		
Group 20	24.59	23.16
Group 30	45.47	39.40
Group 40	35.99	43.78
Group 50	57.55	55.47

Note: Outcomes of the volatility modeling of the Ibovespa by applying an APARCH model, as described in Equation 4.1:
$r_t = \mu + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t$
$a_t = \sigma_t \epsilon_t$
$\sigma_t^{\delta} = \alpha_0 + \sum_{i=1}^{s} \alpha_i (|a_{t-i}| - \gamma_i a_{t-1})^{\delta} + \sum_{i=1}^{m} \beta_i \sigma_{t-i}^{\delta}$
The variance model was estimated with orders 1 and 1 and a Student-t distribution to the daily series of (nominal and dividend-adjusted) log-returns of the index from 2005 to 2018. We present versions with and without autoregressive and moving average components in the ARMA model, but both without the constant term. The estimates are shown with an indication of the statistical significance, considering robust standard errors. ***: significant at 1%; **: significant at 5%; *: significant at 10%.

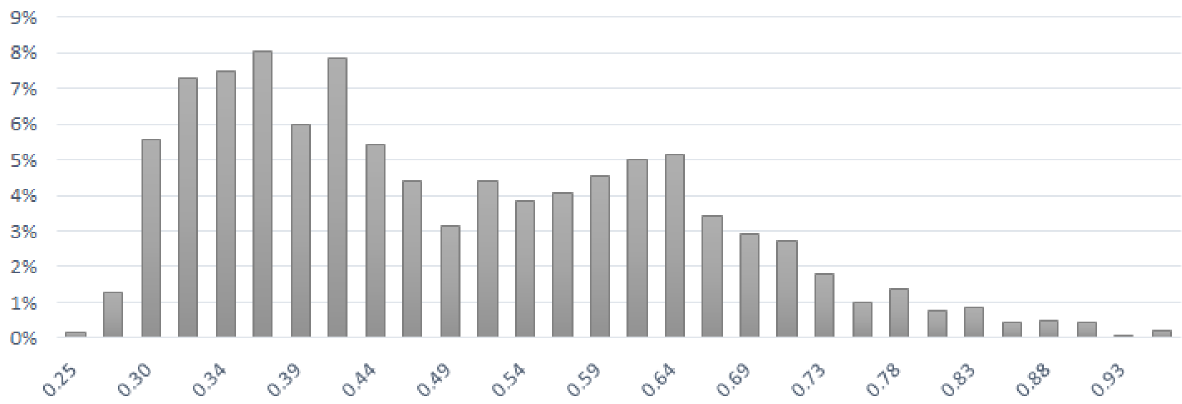

Figure 4.2: Frequency distribution of the asymmetry parameter of the Ibovespa daily volatility

Note: Asymmetry measured by the parameter γ of an APARCH(1,1)ARMA(1,1) model, t-distributed, estimated in moving windows of 2 thousand days, based on the Ibovespa returns from 2005 to 2018.

4.3.2 Outcomes of the Relationship between Asymmetry and Attention

Before regressing the $gamma_t$ series on the explaining variables, we winsorized these asymmetry indices at 5% and 95%. Each observation of $gamma_t$ was then associated with the observations of the other variables that correspond to the last day of the moving windows. In other words, $gamma_1$, estimated from r_1 to r_{2000}, was associated with GSV_{2000}, VM_{2000} and so forth.

Table 4.5 describes the outcomes of the regression of the asymmetry on explaining variables of the model presented in Equation 4.2. To verify whether the results persist in different configurations, we combined scenarios with and without control variables and with and without orders in the ARMA model. Besides that, the data were submitted both to a Fixed Effects Linear Model (FELM) and to a GLS model. The fixed effects and the clustering of errors were designed using year dummies, which was shown to be highly significant.

The results corroborate with our main research hypothesis (H_1). In all the cases, the investor attention measured by the Google search volume ($\log GSV_t$) showed a negative and statistically significant effect in the Ibovespa asymmetry. This is an evidence of the ostrich effect, proposed by Karlsson et al. (2009), in the temporal fluctuation of the index volatility asymmetry. The results are very similar in the specifications with and without orders in the ARMA model.

All risk factors presented relevant and statistically significant coefficients. Size and *momentum* presented negative estimates, while price-to-book showed positive values. This means that the volatility asymmetry of the index is higher in moments of lower market capitalization, after a downward trend of the index and when the market value

Table 4.3: Parameters estimation of the longitudinal determinants of Ibovespa daily volatility asymmetry

	γ_t from APARCH(1,1)ARMA(1,1)				γ_t from APARCH(1,1)ARMA(0,0)			
	FELM	GLS	FELM	GLS	FELM	GLS	FELM	GLS
$\log GSV_t$	-0.02^{**}	-0.02^{**}	-0.02^{***}	-0.02^{**}	-0.03^{**}	-0.03^{***}	-0.03^{***}	-0.03^{***}
	(0.04)	(0.02)	(0.00)	(0.03)	(0.05)	(0.00)	(0.00)	(0.00)
$\log Size_t$			-1.43^{*}	-1.43^{***}			-1.69^{**}	-1.69^{***}
			(0.08)	(0.00)			(0.04)	(0.00)
Mom_t			-0.11^{*}	-0.11^{***}			-0.16^{***}	-0.16^{***}
			(0.07)	(0.00)			(0.00)	(0.00)
$P2B_t$			0.32^{*}	0.32^{***}			0.46^{**}	0.46^{***}
			(0.06)	(0.00)			(0.03)	(0.00)
$D2E_t$			0.00	0.00			-0.00	-0.00^{*}
			(0.76)	(0.29)			(0.53)	(0.07)
DoF	1,397	1,404	1,393	1,404	1,445	1,452	1,441	1,452
Residuals								
Std. errors	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07
Min.	-0.20	-0.21	-0.39	-0.38	-0.15	-0.14	-0.42	-0.44
Q1	-0.05	-0.05	-0.04	-0.04	-0.05	-0.04	-0.04	-0.05
Med.	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
Q3	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Max.	0.41	0.42	0.42	0.40	0.33	0.31	0.31	0.32

Note: Estimates of the regression of the volatility asymmetry index of the Ibovespa on a measure of investor attention (Google search volume), risk factors (size, *momentum* and price-to-book ratio) and the leverage ratio of the index, as stated in Equation 4.2:
$\gamma_t = c + \psi \log GSV_t + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t$
The asymmetry was measured by the specific coefficient of aa APARCH(1,1) model. We present the outcomes for versions with and without orders in the ARMA model, with and without control variables, and by performing a linear modeling with year fixed effects (FELM) and a generalized least squares modeling (GLS), also with year fixed effects. Standard errors are clustered by year. Estimates of the constant and fixed effects coefficients are omitted. The variance model was built with a t-distributed random variable using daily series of (nominal and dividend-adjusted) log-returns of the index from 2005 to 2018. Estimates are presented followed by their p-values in brackets. ***: significant at 1%; **: significant at 5%; *: significant at 10%.

is large compared to the book value of the stocks that comprise it. On the other hand, we did not find evidences that leverage of the firms influences the asymmetry fluctuation of the index. The outcomes do not indicate the existence of the leverage effect described by Black (1976) and Christie (1982). This effect, however, might be clearer in tests with lower frequencies, given that leverage drifts in daily series tend to be more smooth and temporary (Avramov et al., 2006). All these results for the control variables are in line with the cross-sectional investigation of the US market made by Dzieliński et al. (2018).

Table 4.4 presents additional results, in line with the effects evidenced in the regressions. We subdivided the series in quartiles according to the returns and attention levels and verified the average asymmetry in each group. It is 12% lower in times of high attention when compared to times of low attention (0.54 vs. 0.47). This difference is statistically significant, with a p-value of the t-test below 1%. The percentage difference is even higher in bad times (days with steeper market falls, reflected in the quartile with more negative returns), reaching 16% (0.58 vs. 0.49). Low attention quartiles show average asymmetry systematically higher than high attention ones, even comparing intermediary ones, although in some cases the difference is not perceived with two decimal places. This monotonically decreasing pattern does not occur only in the quartile of more positive returns, but in this case days with highest attention are still less asymmetrical, on average, compared to days with lowest attention.

Table 4.4: Average asymmetry in each quartile of returns and attention

		Negative returns Q1	Q2	Positive returns Q3	Q4	Whole sample
Lower attention	Q1	0.58	0.57	0.49	0.50	0.54
	Q2	0.54	0.51	0.49	0.51	0.51
Higher attention	Q3	0.49	0.47	0.46	0.47	0.47
	Q4	0.49	0.47	0.45	0.49	0.47
Whole sample		0.51	0.49	0.47	0.49	0.49

Note: Averages of the asymmetry index in each sample subset. Daily observations were divided in quartiles according to attention and return levels. We used Ibovespa daily data from 2013 to 2018. Attention was measure by Google search volume of the keyword "bovespa". The asymmetry corresponds to the parameter γ of an APARCH(1,1) model with unitary orders in the ARMA model of the expected values.

The relationship we found between asymmetry and attention differs from the one presented by Dzieliński et al. (2018) in the US market. The authors found a positive correlation between attention and asymmetry. It is important to highlight that, different from our work, they used cross-sectional data, on a monthly basis, and at the firm level. Besides that, and perhaps more relevant, attention Dzieliński et al. (2018) used the number of analysts following each firm as a proxy of attention. This measure captures a type of attention that is distinct from the one on which we focus. As analyzed in previous studies

(Da et al., 2011; Dimpfl & Jank, 2016; Kita & Wang, 2012), since it is a popular, free and non-customized tool, Google search engine is used substantially to subsidize the decision making of retail investors. The behavior of this class of investors is more susceptible to biases and is commonly associated with to the one of noise traders, inducing a non-fundamental volatility to the markets. This phenomenon, associated with the ostrich effect, leads to the asymmetry reduction evidenced in moments of higher attention.

4.3.3 Isolating the Effects of Professional and Retail Attention on Asymmetry

In this section we perform an additional analysis of the impact of the attention on the volatility asymmetry. The purpose is to verify whether, in fact, the non-professional investor attention is the one that has a behavior of asymmetry reducer.

We obtained the volume of data searched and news read at Bloomberg (BSV), to perform this analysis. Besides being traditionally the more adopted information provider by analysts and professional traders, Bloomberg charges a monthly or annual fee and requires specific terminals to provide the access, offering several components and services (Ben-Rephael et al., 2017).

As described in the section 4.2.1, the variable "NEWS HEAT READ DMAX" of this financial information provider is an index that varies from 0 to 4 and combines the quantity of accesses to news that cite the stocks with active ticker search queries for related news. We assigned 0 to fields with unavailable information. As it is available only for individual stocks, we performed a principal component analysis (PCA) of the (correlated) series of the five most representative tickers that comprise the Ibovespa in September 2019. Four explaining variables resulted from this analysis: $BSV1_t$ a $BSV4_t$. They are linearly not correlated and are proxies of how attentive professional investors are with respect to the Ibovespa.

We tested Equation 4.3 to verify whether professional attention affects the Ibovespa asymmetry.

$$\begin{aligned}\gamma_t = \kappa_0 + \kappa_1 BSV1_t + \kappa_2 BSV2_t + \kappa_3 BSV3_t + \kappa_4 BSV4_t + \\ + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t\end{aligned} \tag{4.3}$$

After that, we regressed the Google search volume on the variables that represent Bloomberg search volume and obtained the residuals. They capture the retail investor attention ($rGSV_t$) whose variance is not explained by the professional attention. We then regress the volatility asymmetry on these residuals to verify the effect induced in asymmetry coming solely from non-sophisticated investor attention. Equations 4.4 and 4.5 represent this modeling. The equation to obtain $rGSV_t$ is similar to the general form presented in Chapter 1 (Equation 1.2), with the difference that here we have four professional attention variables as regressors.

$$GSV_t = \chi_0 + \chi_1 BSV1_t + \chi_2 BSV2_t + \chi_3 BSV3_t + \chi_4 BSV4_t + rGSV_t \tag{4.4}$$

$$\gamma_t = c + \nu rGSV_t + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t \tag{4.5}$$

Just as in the previous regressions, we used robust standard errors clustered by year. We adopted year fixed effects linear models (FELM) and generalized least squares models (GLS). The same control variables were used (the risk factors size, price-to-book and *momentum*, as well as the leverage ratio).

The results do not allow rejecting the conjecture that there is no asymmetry effect arising from professional attention variations. The variables that represent Bloomberg searches did not present statistical significance in any configuration. Nonetheless, the non-professional attention present a negative and significant coefficient in all setups. In line with the previous regressions, the control variables presented significant effects, except for the leverage ratio. We found evidences that converge with our second hypothesis (H_2) that it is the retail attention that mitigates the countercyclical behavior of the volatility. This result is in line with the higher vulnerability of this investor class to the ostrich effect (Karlsson et al., 2009) and with the hypothesis of price pressure due to non-professional excess attention (Barber & Odean, 2008).

4.4 Final Considerations of the Chapter

We verified in this work the investor attention effect on the volatility asymmetry of the stock market. Based on previous studies that evidenced relationships between attention and volatility (Dimpfl & Jank, 2016; Tantaopas et al., 2016) and between attention and returns (ostrich effect, of Karlsson et al. (2009)), we hypothesize that attention mitigates volatility asymmetry levels.

We used daily market data of Brazilian stocks for this analysis. The asymmetry was obtained by the specific parameter of an APARCH model, which is part of a family that is widely used to represent the market volatility. Attention was measured by the search volume performed at Google, a metric with frequency and properties very appropriate for studies of this kind. Variables such as those ones have increasingly become important given digital inclusion, the volume of information generated over the Internet and the propagation of online services.

The results confirmed our conjecture that the attention, particularly the non-

Table 4.5: Estimates of the effect of professional and retail attention on the asymmetry of the Ibovespa daily volatility

	$\gamma_t \sim$ Professional attention				$\gamma_t \sim$ Retail attention			
	FELM	GLS	FELM	GLS	FELM	GLS	FELM	GLS
$BSV1_t$	-0.00	-0.00	-0.00	-0.00				
	(0.56)	(0.21)	(0.81)	(0.70)				
$BSV2_t$	-0.00	-0.00	-0.00	-0.00				
	(0.47)	(0.69)	(0.21)	(0.55)				
$BSV3_t$	-0.00	-0.00	-0.00	-0.00				
	(0.97)	(0.96)	(0.86)	(0.84)				
$BSV4_t$	-0.00	-0.00	-0.00	-0.00				
	(0.67)	(0.58)	(0.90)	(0.86)				
$rGSV_t$					-0.03^{**}	-0.03^{***}	-0.02^{**}	-0.02^{**}
					(0.02)	(0.01)	(0.01)	(0.03)
$\log Size_t$			-1.42^{*}	-1.42^{***}			-1.43^{*}	-1.43^{*}
			(0.08)	(0.00)			(0.08)	(0.00)
Mom_t			-0.11^{*}	-0.11^{***}			-0.11^{*}	-0.11^{***}
			(0.07)	(0.00)			(0.07)	(0.00)
$P2B_t$			0.31^{*}	0.31^{***}			0.31^{*}	0.31^{***}
			(0.07)	(0.00)			(0.06)	(0.00)
$D2E_t$			0.00	0.00			0.00	-0.00^{*}
			(0.77)	(0.32)			(0.76)	(0.29)
DoF	1,394	1,404	1,390	1,404	1,397	1,404	1,393	1,404
Residuals								
Std. errors	0.08	0.08	0.07	0.07	0.08	0.08	0.07	0.07
Min.	-0.20	-0.21	-0.40	-0.38	-0.20	-0.14	-0.39	-0.37
Q1	-0.05	-0.05	-0.04	-0.04	-0.05	-0.04	-0.04	-0.04
Med.	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
Q3	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Max.	0.41	0.42	0.42	0.41	0.41	0.31	0.42	0.40

Note: The four columns on the left show the outcomes of the regression of the Ibovespa volatility asymmetry on professional attention measures, represented on Equation 4.3:
$\gamma_t = \kappa_0 + \kappa_1 BSV1_t + \kappa_2 BSV2_t + \kappa_3 BSV3_t + \kappa_4 BSV4_t + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t$
These measures ($BSV1_t$ a $BSV4_t$) result from a principal component analysis performed in series of Bloomberg search volume of the most representative stocks that comprise the Ibovespa. The fours columns on the right present the outcomes of the regression of the asymmetry on the retail investor attention $rGSV_t$, which is the investor attention not explained by professional attention, as it is shown on Equations 4.4 and 4.5:
$GSV_t = \chi_0 + \chi_1 BSV1_t + \chi_2 BSV2_t + \chi_3 BSV3_t + \chi_4 BSV4_t + rGSV_t$
$\gamma_t = c + \nu rGSV_t + \upsilon \log Size_t + \lambda Mom_t + \zeta P2B_t + \rho D2E_t + \eta_t + \epsilon_t$
In both scenarios, we present versions with and without size, *momentum*, price-to-book and leverage variables. The Ibovespa asymmetry over the period was measured by the specific coefficient of a t-distributed APARCH(1,1) and ARMA(1,1) model of the daily series of (nominal and dividend-adjusted) log-returns, using moving windows of 2 thousand days that cover the period from 2005 to 2018. Linear models with year fixed effects (FELM) and generalized least squares models (GLS) were performed, with standard errors clustered by year. Estimates are followed but their p-value. ***: significant at 1%; **: significant at 5%; *: significant at 10%.

professional one, reduces the intensity of an stylized fact of financial markets: volatility is higher in bad times than in equivalent good times. When this investor "hides his head inside a hole", his role of volatility inducer becomes less expressive, reducing the discrepancy that volatility presents in good and bad times. We found that moments with less attention record volatility asymmetry on average 12% lower than moments with high attention. Our results are robust to different setups, to the inclusion of control variables and evidence that it is the retail attention that causes this impact on asymmetry.

Our findings offer some contributions to the literature and to practitioners. Firstly, we did not find an analysis of the effect of attention on daily asymmetry in the stock market. Neither did we find an study of the asymmetry focusing in retail investor attention, usually with behavioral biases that are more notable, and that challenge classical theories that presuppose the full rationality of the agents. This innovation resulted in outcomes that are different from the ones found by Dzieliński et al. (2018) in their analysis of the impact of analyst coverage on the US stock market, although none of the studies found significant evidence of the classical leverage effect reported by Black (1976) and Christie (1982). This pattern is in line with the idea of Avramov et al. (2006) that the transitory and smooth behavior of daily changes in leverage limits the identification of the leverage effect. Studies that help to understand how our cognitive resources influence the market behavior support decisions related to risk management, asset pricing and information releases.

Future studies may verify the effect of the attention on risk premiums, given its influence on levels and asymmetry of risk. Besides, several research lines can be developed to better understand the determinants and consequences of the ostrich effect in financial markets. A promising alternative would be to compare the proportions of this effect in emerging and developed markets, given their differences with respect to maturity, stability, concentration, share of non-professional investors, among other factors.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43–76 (cit. on p. 76).

Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *The Review of Financial Studies*, *28*(1), 33–72 (cit. on pp. 75, 81).

Avramov, D., Chordia, T., & Goyal, A. (2006). The impact of trades on daily volatility. *The Review of Financial Studies*, *19*(4), 1241–1277 (cit. on pp. 74, 96, 100).

Backus, D. K., Routledge, B. R., & Zin, S. E. (2004). Exotic preferences for macroeconomists. *NBER Macroeconomics Annual*, *19*, 319–390 (cit. on p. 82).

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of german stocks. *Financial Markets and Portfolio Management*, *25*(3), 239 (cit. on p. 87).

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *The Review of Financial Studies*, *21*(2), 785–818 (cit. on pp. 10, 20, 25, 28, 43, 47, 50, 51, 68, 86, 98).

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics*, *116*(1), 1–53 (cit. on p. 82).

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *The Review of Financial Studies*, *13*(1), 1–42 (cit. on pp. 77, 78, 81).

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*(1), 1–27 (cit. on p. 85).

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009–3047 (cit. on pp. 10, 15, 43, 88, 97).

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, 177–181 (cit. on pp. 73, 77, 86, 91, 96, 100).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327 (cit. on pp. 60, 90).

Breen, W., Glosten, L. R., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of Finance*, *44*(5), 1177–1189 (cit. on p. 74).

Campbell, J. Y., & Hentschel, L. (1991). *No news is good news: An asymmetric model of changing volatility in stock returns* (tech. rep.). National Bureau of Economic Research. (Cit. on pp. 73, 74, 77, 89).

Cheung, Y.-W., & Ng, L. K. (1992). Stock price dynamics and firm size: An empirical investigation. *The Journal of Finance*, *47*(5), 1985–1997 (cit. on p. 81).

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, *10*(4), 407–432 (cit. on pp. 77, 79, 91, 96, 100).

Chronopoulos, D. K., Papadimitriou, F. I., & Vlastakis, N. (2018). Information demand and stock return predictability. *Journal of International Money and Finance*, *80*, 59–74 (cit. on p. 87).

Cochrane, J. H. (2001). *Asset pricing*. Princeton University Press. (Cit. on p. 74).

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499 (cit. on pp. 10, 11, 20, 43, 47, 48, 52, 53, 56, 57, 60, 65, 66, 68, 75, 87, 97).

Da, Z., Gurun, U. G., & Warachka, M. (2014). Frog in the pan: Continuous information and momentum. *The Review of Financial Studies*, *27*(7), 2171–2218 (cit. on p. 82).

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, *22*(2), 171–192 (cit. on pp. 10, 47, 48, 54, 58, 59, 65, 66, 68, 75, 86–88, 97, 98).

Ding, Z., Granger, C., & Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*, 83–106 (cit. on p. 89).

Dzieliński, M., Rieger, M. O., & Talpsepp, T. (2018). Asymmetric attention and volatility asymmetry. *Journal of Empirical Finance*, *45*, 59–67 (cit. on pp. 75, 89, 91, 92, 96, 100).

Engelberg, J., & Parsons, C. A. (2016). Worrying about the stock market: Evidence from hospital admissions. *The Journal of Finance*, *71*(3), 1227–1250 (cit. on p. 75).

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimate of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*, 987–1001 (cit. on p. 90).

Feng, L., & Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, *9*(3), 305–351 (cit. on p. 88).

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual investors and volatility. *The Journal of Finance*, *66*(4), 1369–1406 (cit. on p. 88).

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*(1), 3 (cit. on pp. 73, 77).

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300 (cit. on p. 82).

Gasparini, D. C. d. S. Et al. (2013). Verificação da performance de modelos aparch assimétricos aplicados a dados financeiros (cit. on p. 90).

Geweke, J. (1986). Comment on: Modelling the persistence of conditional variances. *Econometric Reviews*, *5*, 57–61 (cit. on p. 90).

Gherzi, S., Egan, D., Stewart, N., Haisley, E., & Ayton, P. (2014). The meerkat effect: Personality and market returns affect investors' portfolio monitoring behaviour. *Journal of Economic Behavior & Organization*, *107*, 512–526 (cit. on p. 75).

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801 (cit. on pp. 77, 90).

Griffin, J. M., Nardari, F., & Stulz, R. M. (2004). *Stock market trading and market conditions* (tech. rep.). National Bureau of Economic Research. (Cit. on p. 85).

Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667–686 (cit. on p. 85).

Higgins, M. L., & Bera, A. K. (1992). A class of nonlinear arch models. *International Economic Review*, 137–158 (cit. on p. 90).

Jacobs, H., & Weber, M. (2012). The trading volume impact of local bias: Evidence from a natural experiment. *Review of Finance*, *16*(4), 867–901 (cit. on p. 87).

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, *38*(2), 95–115 (cit. on pp. 75, 82, 85, 86, 94, 98).

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics*, *104*(2), 321–338 (cit. on p. 75).

Kita, A., & Wang, Q. (2012). Investor attention and fx market volatility. *Available at SSRN 2022100* (cit. on pp. 87, 97).

Lehmann, B. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, (105), 1–28 (cit. on p. 74).

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282 (cit. on p. 85).

Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, *397*(6719), 498 (cit. on pp. 47, 49, 50, 59, 70, 86, 88).

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370 (cit. on p. 77).

Patton, A. J., & Timmermann, A. (2008). Generalized forecast errors, a change of measure, and forecast optimality. Citeseer. (Cit. on p. 82).

Pentula, S. (1986). Comments on “modelling the persistence of conditional variances.” *Econometric Reviews*, *5*, 71–74 (cit. on p. 90).

Pindyck, R. S. (1984). Risk, inflation, and the stock market. *American Economic Review*, *74*, 334–351 (cit. on pp. 73, 77).

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: Basic concepts and examples. *Mixed-Effects Models in S and S-Plus*, 3–56 (cit. on p. 91).

Schneider, S. L. (2001). In search of realistic optimism: Meaning, knowledge, and warm fuzziness. *American Psychologist*, *56*(3), 250 (cit. on p. 84).

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, *44*(5), 1115–1153 (cit. on pp. 73, 77).

Schwert, G. W. (1990). Stock volatility and the crash of’87. *The Review of Financial Studies*, *3*(1), 77–102 (cit. on p. 90).

Sims, C. A. (1984). Martingale-like behavior of prices and interest rates, (Discussion Paper No. 205) (cit. on p. 74).

Talpsepp, T., & Rieger, M. O. (2010). Explaining asymmetric volatility around the world. *Journal of Empirical Finance*, *17*(5), 938–956 (cit. on pp. 75, 89).

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in asia-pacific stock markets. *Pacific-Basin Finance Journal*, *38*, 107–124 (cit. on pp. 29, 30, 32, 43, 51, 57, 58, 98).

Tauchen, G., Zhang, H., & Liu, M. (1996). Volume, volatility, and leverage: A dynamic analysis. *Journal of Econometrics*, *74*(1), 177–208 (cit. on p. 76).

Taylor, S. (1986). *Modelling financial time series*. Wiley, New York. (Cit. on p. 90).

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics*, *25*(1), 3–22 (cit. on p. 77).

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291 (cit. on p. 74).

Van Nieuwerburgh, S., & Veldkamp, L. (2006). Learning asymmetries in real business cycles. *Journal of Monetary Economics*, *53*(4), 753–772 (cit. on p. 82).

Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, *36*(6), 1808–1821 (cit. on p. 87).

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), 931–955 (cit. on p. 90).

Zeelenberg, M., Van Dijk, W. W., Manstead, A. S., & vanr de Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition & Emotion*, *14*(4), 521–541 (cit. on p. 85).

5 Concluding Remarks

Give whatever you are doing and
whoever you are with the gift of your
attention.

Jim Rohn (1930-2009)
Entrepreneur, author and
motivational speaker

Over the last years, collaborations between the theory of behavioral economics and methodologies of the cognitive science allowed the access of novel data from a variety of new sources — the human brain, conference transcripts, genetic information, online activities, among others —, which revealed new findings about cognitive processes that influenced the financial decision-making. This has challenged evidences and theories based on classical models of information processing.

These findings are taken into account in the development of financial products, in conferences and statements about the performance of firms, and even in the implementation of nudges — indirect suggestions and positive reinforcement to influence the behavior of groups.

The increasing generation and release of information motivate the study of attention, one of the most relevant cognitive processes. It is practically impossible to pay attention to everything, and things that capture more attention naturally tend to impact more the financial decisions and consequently the asset prices. Attention is a scarce resource, with finite capacity and influenced by saliences in some attributes — we started from this general problem to develop the essays reported in this study.

More precisely, we aimed to better understand the fluctuations in attention levels in the financial market and investigate its impact in specific variables. Particularly, we verified the effects of attention variations on the volatility of prices, on the asymmetry of that volatility, and on the efficiency (in terms of predictability level) of the Brazilian stock market.

The study went through some stages: firstly, the conception of the ideas and assessment of their innovation potential, then the analysis of previous related papers, and, finally, the design and implementation of adequate empirical methods that allowed an evaluation of the hypotheses with a reasonable confidence level.

We found a significant volatility induction in moments of higher attention, combined with a reduction in the asymmetry of this volatility. Besides that, we verified that the market becomes less predictable when investors, particularly the professional ones, are more attentive. Above all, our findings showed the potential that variables arising from the increasing use of the Internet may have in research related to behavioral finance. In general, these measures are easily accessible, have higher frequencies than their alternatives and reasonable customization levels.

Our work complements several previous studies, both in other emergent and in more mature markets. Our hypothesis tests allowed a comprehensive analysis of the Brazilian stock market and the results were in line with conjectures such as the price pressure induced by excess of attention, the information discovery hypothesis and the ostrich effect.

However, we innovated by studying for the first time (as far as we searched) the attention effects on the daily volatility of the most important Brazilian stock index, as well as on the asymmetry of this volatility. Besides, this study was a pioneer in isolating the attention of the retail investor from the attention of the professional one in tests of market efficiency and volatility asymmetry.

We believe this document opens promising venues of research, with the potential to contribute in a relevant way to the theoretical framework and to practitioners. An alternative is the replication of the hypothesis tests in other markets (monetary, ForEx, commodities, cryptocurrencies, among others) or comparative studies that analyze differences between emergent and more mature markets. Other studies may assess user activity data of social media, such as Twitter or LinkedIn, for more specialized analysis. The development of investment strategies, or trading algorithms, that incorporate information about investor attention may have their performance improved.

Research proposals to better understand the determinants of attention fluctuations of different investor classes are also viable. This type of approach can be supported by cognitive and behavioral science tools to verify mechanisms through which attention exerts influence in financial markets. One can also deepen the analysis of the different effects presented in this study, identifying situations when there is a dominance of one of them, as well as the aspects that determine these dynamics.

To finalize, event studies allow evidencing differences in the fluctuation patterns of attention during critical moments, such as public offerings, presidential elections and corporate corruption scandals. These approaches may require intraday frequencies, but comparative studies of those events can result in finding effects of large magnitude and a decisive role of the attention.

References

Adachi, Y., Masuda, M., & Takeda, F. (2017). Google search intensity and its relationship to the returns and liquidity of japanese startup stocks. *Pacific-Basin Finance Journal*, *46*, 243–257 (cit. on p. 31).

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of financial markets*, *5*(1), 31–56 (cit. on pp. 19, 25).

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, *61*(1), 43–76 (cit. on p. 76).

Andreassen, P. B., & Kraus, S. J. (1990). Judgmental extrapolation and the salience of change. *Journal of Forecasting*, *9*(4), 347–372 (cit. on p. 10).

Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *The Review of Financial Studies*, *28*(1), 33–72 (cit. on pp. 75, 81).

Avramov, D., Chordia, T., & Goyal, A. (2006). The impact of trades on daily volatility. *The Review of Financial Studies*, *19*(4), 1241–1277 (cit. on pp. 74, 96, 100).

B3. (2015). Metodologia do índice bovespa (cit. on p. 31).

Backus, D. K., Routledge, B. R., & Zin, S. E. (2004). Exotic preferences for macroeconomists. *NBER Macroeconomics Annual*, *19*, 319–390 (cit. on p. 82).

Bank, M., Larch, M., & Peter, G. (2011). Google search volume and its influence on liquidity and returns of german stocks. *Financial Markets and Portfolio Management*, *25*(3), 239 (cit. on p. 87).

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional in. *The Review of Financial Studies*, *21*(2), 785–818 (cit. on pp. 10, 20, 25, 28, 43, 47, 50, 51, 68, 86, 98).

Barberis, N., Huang, M., & Santos, T. (2001). Prospect theory and asset prices. *The Quarterly Journal of Economics*, *116*(1), 1–53 (cit. on p. 82).

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307–343 (cit. on p. 19).

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *The Review of Financial Studies*, *13*(1), 1–42 (cit. on pp. 77, 78, 81).

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*(1), 1–27 (cit. on p. 85).

Benartzi, S., & Thaler, R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics*, *110*(1), 73–92 (cit. on p. 48).

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009–3047 (cit. on pp. 10, 15, 43, 88, 97).

Bijl, L., Kringhaug, G., Molnár, P., & Sandvik, E. (2016). Google searches and stock returns. *International Review of Financial Analysis*, *45*, 150–156 (cit. on p. 53).

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*, 177–181 (cit. on pp. 73, 77, 86, 91, 96, 100).

Black, F. (1986). Noise. *The Journal of Finance*, *41*(3), 528–543 (cit. on p. 48).

BMFBovespa. (2017). Participação dos investidores. Available at bmfbovespa.com.br. (Cit. on p. 60).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327 (cit. on pp. 60, 90).

Breen, W., Glosten, L. R., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of Finance*, *44*(5), 1177–1189 (cit. on p. 74).

Brooks, C. (1998). Predicting stock index volatility: Can market volume help? *Journal of Forecasting*, *17*(1), 59–80 (cit. on pp. 47, 68).

Burton, E., & Shah, S. (2013). *Behavioral finance: Understanding the social, cognitive, and economic debates*. John Wiley & Sons. (Cit. on pp. 10, 48).

Campbell, J. Y., & Hentschel, L. (1991). *No news is good news: An asymmetric model of changing volatility in stock returns* (tech. rep.). National Bureau of Economic Research. (Cit. on pp. 73, 74, 77, 89).

Chan, W. S. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, *70*(2), 223–260 (cit. on p. 25).

Cheung, Y.-W., & Ng, L. K. (1992). Stock price dynamics and firm size: An empirical investigation. *The Journal of Finance*, *47*(5), 1985–1997 (cit. on p. 81).

Chiang, T. C., & Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, *34*(8), 1911–1921 (cit. on p. 41).

Choi, H., & Varian, H. (2009). Predicting initial claims for unemployment benefits. *Google Inc*, 1–5 (cit. on p. 11).

Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage and interest rate effects. *Journal of Financial Economics*, *10*(4), 407–432 (cit. on pp. 77, 79, 91, 96, 100).

Chronopoulos, D. K., Papadimitriou, F. I., & Vlastakis, N. (2018). Information demand and stock return predictability. *Journal of International Money and Finance*, *80*, 59–74 (cit. on p. 87).

Cochrane, J. H. (2001). *Asset pricing*. Princeton University Press. (Cit. on p. 74).

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance*, *66*(5), 1461–1499 (cit. on pp. 10, 11, 20, 43, 47, 48, 52, 53, 56, 57, 60, 65, 66, 68, 75, 87, 97).

Da, Z., Gurun, U. G., & Warachka, M. (2014). Frog in the pan: Continuous information and momentum. *The Review of Financial Studies*, *27*(7), 2171–2218 (cit. on p. 82).

Damien, C., Ahmed, B. H. A. Et al. (2013). *Predicting financial markets with google trends and not so random keywords* (tech. rep.). (Cit. on p. 47).

Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance*, *52*(3), 1035–1058 (cit. on p. 25).

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703–738 (cit. on p. 10).

De Souza, H. E., Barbedo, C. H. D. S., & Araújo, G. S. (2017). Does investor attention affect trading volume in the brazilian stock market? *Research in International Business and Finance* (cit. on p. 56).

de Ávila Marques, T., Machado Guimarães, T., & Macial Peixoto, F. (2015). A concentração acionária no brasil: Análise dos impactos no desempenho, valor e risco das empresas. *RAM. Revista de Administração Mackenzie*, *16*(4) (cit. on p. 58).

Dimpfl, T., & Jank, S. (2016). Can internet search queries help to predict stock market volatility? *European Financial Management*, *22*(2), 171–192 (cit. on pp. 10, 47, 48, 54, 58, 59, 65, 66, 68, 75, 86–88, 97, 98).

Ding, Z., Granger, C., & Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*, 83–106 (cit. on p. 89).

Dzieliński, M., Rieger, M. O., & Talpsepp, T. (2018). Asymmetric attention and volatility asymmetry. *Journal of Empirical Finance*, *45*, 59–67 (cit. on pp. 75, 89, 91, 92, 96, 100).

Eldridge, R. (2016). How social media is shaping financial services. (Cit. on p. 70).

Engelberg, J., & Parsons, C. A. (2016). Worrying about the stock market: Evidence from hospital admissions. *The Journal of Finance*, *71*(3), 1227–1250 (cit. on p. 75).

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimate of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, *50*, 987–1001 (cit. on p. 90).

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007 (cit. on p. 60).

Falkinger, J. (2008). Limited attention as a scarce resource in information-rich economies. *The Economic Journal*, *118*(532), 1596–1620 (cit. on pp. 19, 26, 27).

Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, *38*(1), 34–105 (cit. on p. 32).

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, *64*(5), 2023–2052 (cit. on pp. 20, 24–26).

Feng, L., & Seasholes, M. S. (2005). Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance*, *9*(3), 305–351 (cit. on p. 88).

Foucault, T., Sraer, D., & Thesmar, D. J. (2011). Individual investors and volatility. *The Journal of Finance*, *66*(4), 1369–1406 (cit. on p. 88).

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, *61*(4), 2017–2046 (cit. on p. 49).

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics*, *19*(1), 3 (cit. on pp. 73, 77).

Gabaix, X., & Laibson, D. (2005). Bounded rationality and directed cognition. *Harvard University* (cit. on p. 20).

Garcia, D. (2013). Sentiment during recessions. *The Journal of Finance*, *68*(3), 1267–1300 (cit. on p. 82).

Gasparini, D. C. d. S. Et al. (2013). Verificação da performance de modelos aparch assimétricos aplicados a dados financeiros (cit. on p. 90).

Geweke, J. (1986). Comment on: Modelling the persistence of conditional variances. *Econometric Reviews*, *5*, 57–61 (cit. on p. 90).

Gherzi, S., Egan, D., Stewart, N., Haisley, E., & Ayton, P. (2014). The meerkat effect: Personality and market returns affect investors' portfolio monitoring behaviour. *Journal of Economic Behavior & Organization*, *107*, 512–526 (cit. on p. 75).

Gillings, M. R., Hilbert, M., & Kemp, D. J. (2016). Information in the biosphere: Biological and digital worlds. *Trends in ecology & evolution*, *31*(3), 180–189 (cit. on p. 9).

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012–1014 (cit. on p. 11).

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779–1801 (cit. on pp. 77, 90).

Gomes, P., & Taamouti, A. (2016). In search of the determinants of european asset market comovements. *International Review of Economics & Finance*, *44*, 103–117 (cit. on p. 10).

Griffin, J. M., Nardari, F., & Stulz, R. M. (2004). *Stock market trading and market conditions* (tech. rep.). National Bureau of Economic Research. (Cit. on p. 85).

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, *70*(3), 393–408 (cit. on pp. 19, 21, 22, 51).

Grullon, G., Kanatas, G., & Weston, J. P. (2004). Advertising, breadth of ownership, and liquidity. *The Review of Financial Studies*, *17*(2), 439–461 (cit. on pp. 20, 24).

Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667–686 (cit. on p. 85).

Higgins, M. L., & Bera, A. K. (1992). A class of nonlinear arch models. *International Economic Review*, 137–158 (cit. on p. 90).

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, *56*(4), 1533–1597 (cit. on pp. 9, 48, 49).

Huddart, S., Lang, M., & Yetman, M. H. (2009). Volume and price patterns around a stock's 52-week highs and lows: Theory and evidence. *Management Science*, *55*(1), 16–31 (cit. on pp. 10, 47, 51).

Institute, B.-T. (2019). Financial market data and analysis global share & segment sizing (cit. on p. 14).

Jacobs, H., & Weber, M. (2012). The trading volume impact of local bias: Evidence from a natural experiment. *Review of Finance*, *16*(4), 867–901 (cit. on p. 87).

Johansson, E. (2014). Creating daily search volume data from weekly and daily data (cit. on pp. 31, 60).

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Prentice-Hall Englewood Cliffs, NJ. (Cit. on pp. 9, 68).

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, *93*(5), 1449–1475 (cit. on p. 41).

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part i* (pp. 99–127). World Scientific. (Cit. on p. 48).

Kaniel, R., Saar, G., & Titman, S. (2008). Individual investor trading and stock returns. *The Journal of Finance*, *63*(1), 273–310 (cit. on p. 10).

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, *38*(2), 95–115 (cit. on pp. 75, 82, 85, 86, 94, 98).

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics*, *104*(2), 321–338 (cit. on p. 75).

Kita, A., & Wang, Q. (2012). Investor attention and fx market volatility. *Available at SSRN 2022100* (cit. on pp. 87, 97).

KPMG. (2012). Social media in financial services. Available at kpmg.de. (Cit. on p. 70).

Kristoufek, L. (2013). Can google trends search queries contribute to risk diversification? *Scientific Reports*, *3* (cit. on p. 11).

Lehmann, B. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, (105), 1–28 (cit. on p. 74).

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282 (cit. on p. 85).

Lux, T., & Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, *397*(6719), 498 (cit. on pp. 47, 49, 50, 59, 70, 86, 88).

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, *42*(3), 483–510 (cit. on pp. 19, 22, 23, 51).

Ministério do Planejamento, D. e. G. (2015). Mercado de capital brasileiro está em 12º no ranking mundial (cit. on pp. 11, 30).

Miragaya, F. (2017). O efeito dos retornos anormais nas buscas por informação dos investidores. *Masters Thesis* (cit. on pp. 56, 57).

Mondria, J., & Wu, T. (2010). The puzzling evolution of the home bias, information processing and financial openness. *Journal of Economic Dynamics and Control*, *34*(5), 875–896 (cit. on p. 20).

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370 (cit. on p. 77).

Patton, A. J., & Timmermann, A. (2008). Generalized forecast errors, a change of measure, and forecast optimality. Citeseer. (Cit. on p. 82).

Pentula, S. (1986). Comments on "modelling the persistence of conditional variances." *Econometric Reviews*, *5*, 71–74 (cit. on p. 90).

Pindyck, R. S. (1984). Risk, inflation, and the stock market. *American Economic Review*, *74*, 334–351 (cit. on pp. 73, 77).

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: Basic concepts and examples. *Mixed-Effects Models in S and S-Plus*, 3–56 (cit. on p. 91).

Preis, T., Moat, H. S., & Stanley, H. E. (2013). Quantifying trading behavior in financial markets using google trends. *Scientific Reports*, *3*, srep01684 (cit. on p. 11).

Preis, T., Reith, D., & Stanley, H. E. (2010). Complex dynamics of our economic life on different scales: Insights from search engine query data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *368*(1933), 5707–5719 (cit. on p. 47).

Rodolfo, E. R., Barbedo, C. H. d. S., & Val, F. d. F. (2017). Relevance of the attention effect in the brazilian stock market. *Latin American Business Review*, *18*(2), 165–187 (cit. on pp. 56, 57).

Schneider, S. L. (2001). In search of realistic optimism: Meaning, knowledge, and warm fuzziness. *American Psychologist*, *56*(3), 250 (cit. on p. 84).

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance*, *44*(5), 1115–1153 (cit. on pp. 73, 77).

Schwert, G. W. (1990). Stock volatility and the crash of'87. *The Review of Financial Studies*, *3*(1), 77–102 (cit. on p. 90).

Sewell, M. (2007). Behavioural finance. *University of Cambridge*, 1–14 (cit. on p. 42).

Sims, C. A. (1984). Martingale-like behavior of prices and interest rates, (Discussion Paper No. 205) (cit. on p. 74).

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, *96*(2), 158–163 (cit. on pp. 19, 24).

Statista. (2014). Popular online search engines in brazil as of may 2014, based on market share (cit. on p. 31).

Talpsepp, T., & Rieger, M. O. (2010). Explaining asymmetric volatility around the world. *Journal of Empirical Finance*, *17*(5), 938–956 (cit. on pp. 75, 89).

Tantaopas, P., Padungsaksawasdi, C., & Treepongkaruna, S. (2016). Attention effect via internet search intensity in asia-pacific stock markets. *Pacific-Basin Finance Journal*, *38*, 107–124 (cit. on pp. 29, 30, 32, 43, 51, 57, 58, 98).

Tauchen, G., Zhang, H., & Liu, M. (1996). Volume, volatility, and leverage: A dynamic analysis. *Journal of Econometrics*, *74*(1), 177–208 (cit. on p. 76).

Taylor, S. (1986). *Modelling financial time series*. Wiley, New York. (Cit. on p. 90).

Times, F. (2018). Bloomberg and reuters lose data share to smaller rivals (cit. on p. 14).

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics*, *25*(1), 3–22 (cit. on p. 77).

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141–162). Springer. (Cit. on p. 10).

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, *47*(2), 263–291 (cit. on p. 74).

Van Nieuwerburgh, S., & Veldkamp, L. (2006). Learning asymmetries in real business cycles. *Journal of Monetary Economics*, *53*(4), 753–772 (cit. on p. 82).

Van Nieuwerburgh, S., & Veldkamp, L. (2007). Information immobility and the home bias puzzle. *NBER Working Paper No. 13366* (cit. on p. 20).

Vlastakis, N., & Markellos, R. N. (2012). Information demand and stock market volatility. *Journal of Banking & Finance*, *36*(6), 1808–1821 (cit. on p. 87).

Vozlyublennaia, N. (2014). Investor attention, index performance, and return predictability. *Journal of Banking & Finance*, *41*, 17–35 (cit. on pp. 28, 29, 32, 43, 51).

Welagedara, V., Deb, S. S., & Singh, H. (2017). Investor attention, analyst recommendation revisions, and stock prices. *Pacific-Basin Finance Journal*, *45*, 211–223 (cit. on p. 14).

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), 931–955 (cit. on p. 90).

Zeelenberg, M., Van Dijk, W. W., Manstead, A. S., & vanr de Pligt, J. (2000). On bad decisions and disconfirmed expectancies: The psychology of regret and disappointment. *Cognition & Emotion*, *14*(4), 521–541 (cit. on p. 85).